

(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

EDITED BY

CARL MURCHISON

Gordon Allport
Harvard University

Arthur G. Bills
University of Cincinnati

E. A. Bott
University of Toronto

Leonard Carmichael
Smithsonian Institution

Herbert S. Conrad
U. S. Office of Education

Elmer K. Culler
University of Rochester

John F. Dashiell
University of North
Carolina

Roy M. Dorcus
University of California
at Los Angeles

Horace B. English
Ohio State University

Franklin Fearing
University of California
at Los Angeles

Frank A. Geldard
University of Virginia

Clarence H. Graham
Columbia University

Ernest R. Hilgard
Stanford University

H. M. Johnson
Tulane University

Heinrich Klüver
University of Chicago

Carney Landis
Columbia University

K. S. Lashley
Harvard University

Norman L. Munn
Bowdoin College

John Paul Nafe
Washington University

Donald G. Paterson
University of Minnesota

Lee Edward Travis
University of Southern
California

Robert C. Tryon
University of California

Michael J. Zigler
Wellesley College

If this space should be unstamped, this is the regular library edition.
But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not
be resold for a period of five years from date of publication.

VOLUME 41

1956

Copyright, 1956, by The Journal Press
Published Quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.

AUTHOR INDEX

Abramson, H. A.	1, 81, 185, 199	Leopold, Julius	61
Albee, George W.	55	Lepley, William M.	419
Andreas, B. G.	107	London, Ivan D.	157
Ayllon, Teodora	163	Lotsof, Edwin J.	427
Bakan, Paul	369	Maslow, A. H.	247
Blank, Leonard	119	Matarazzo, Joseph D.	347
Brodbeck, Arthur J.	23	Matarazzo, Ruth G.	347
Bruell, Jan H.	55	Mintz, N. L.	247, 459
Chansky, Norman M.	3	Moore, Kate Gordon	127
Charen, Sol	273	Moriarty, D.	243
Clarke, Walter V.	379, 405, 413	Nogee, Philip	23
Comalli, Peter E., Jr.	335	Novak, S.	243
Dillon, Donald	61	Payne, R. B.	47
Edgington, S.	271	Rawn, Mose L.	119
Ekman, Gosta	231	Roff, Merrill	47
Flores, Ivan	301	Sampson, Philip B.	135, 151
Freedman, Mervin	315	Sanford, Nevitt	315
Gerall, Arnold A.	107, 135, 151	Saslow, George	347
Goodenough, D.	107	Simpson, James E.	235
Gorin, M. H.	81	Snyder, Charles W.	177
Green, Russell F.	107, 151	Snyder, Fred W.	177
Halpern, L.	255	Sommer, Robert	163
Hirsch, M. W.	11, 81	Spragg, S. D. S.	107, 135, 151
Humm, Doncaster G.	395	Voor, Joseph H.	437
Jarvik, M. E.	11, 81, 185	Watson, George	323
Jones, A. W.	291	Webster, Harold	315
Koltuv, M.	243	Welch, L.	243
Kugelmass, S.	255	Wapner, Seymour	335
Landis, Carney	61	Werner, Heinz	335
Lennard, H.	185	Zeigler, Martin L.	419

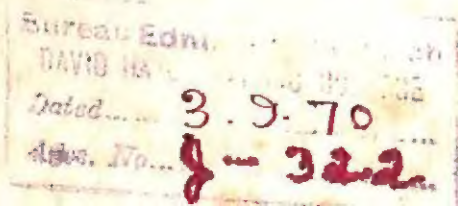


TABLE OF CONTENTS

Threat as a factor in recall in a retroactive paradigm	3
BY NORMAN M. CHANSKY	
Lysergic acid diethylamide (LSD-25): XVIII. Effects of LSD-25 and six related drugs upon handwriting	11
BY M. W. HIRSCH, M. E. JARVIK, AND H. A. ABRAMSON	
Two kinds of conformity: A study of the Riesman typology applied to standards of parental discipline	23
BY ARTHUR J. BRODBECK AND PHILIP NOGEE	
The relation between learning ability scores obtained in different situations	47
BY MERRILL ROFF AND R. B. PAYNE	
A new illusion of apparent movement and the concept of retinal local signs	55
BY JAN H. BRUELL AND GEORGE W. ALBEE	
Changes in flicker-fusion threshold and in choice reaction time induced by electroconvulsive therapy	61
BY CARNEY LANDIS, DONALD DILLON, AND JULIUS LEOPOLD	
Lysergic acid diethylamide (LSD-25): XVII. Tolerance development and its relationship to a theory of psychosis	81
BY H. A. ABRAMSON, M. E. JARVIK, M. H. GORIN, AND M. W. HIRSCH	
Performance levels and transfer effects in compensatory and following tracking as a function of the planes of rotation of control cranks	107
BY R. F. GREEN, D. GOODENOUGH, B. G. ANDREAS, A. A. GERALL, AND S. D. S. SPRAGG	
An experimental method to measure intellectual functioning with verbal and motor factors minimal (1)	119
BY LEONARD BLANK AND MOSE L. RAWN	
Some Arab doctrines on imagination: A chapter in the history of psychology	127
BY KATE GORDON MOORE	
Performance on a tracking task as a function of position, radius, and loading of control cranks: I. Stationary targets	135
BY ARNOLD A. GERALL, PHILIP B. SAMPSON, AND S. D. S. SPRAGG	
Performance on a tracking task as a function of position, radius, and loading of control cranks: II. Moving targets	151
BY ARNOLD A. GERALL, P. B. SAMPSON, RUSSELL F. GREEN, AND S. D. S. SPRAGG	
Derivation of a generalized class of ideal retention curves based on a concept of equipotentiality	157
BY IVAN D. LONDON	
Autism, emphasis, and figure-ground perception	163
BY TEODORO AYLLON AND ROBERT SOMMER	
The effects of monetary reward and punishment on auditory perception	177
BY FRED W. SNYDER AND CHARLES W. SNYDER	
Lysergic acid diethylamide (LSD-25): XII. A preliminary statement of its effects upon interpersonal communication	185
BY H. LENNARD, M. E. JARVIK, AND H. A. ABRAMSON	
Lysergic acid diethylamide (LSD-25) as an adjunct to brief psychotherapy, with special reference to ego enhancement	199
BY HAROLD A. ABRAMSON	
A note on the normal probability function in color vision	231
BY GOSTA EKMEN	

The influence of auditory stimulation on aniseikonic perception: I. A preliminary study	235
BY JAMES E. SIMPSON	
The effect of increased trials on the stability of the <i>PGR</i> in an avoidance situation	243
BY S. NOVAK, D. MORIARTY, M. KOLTUV, AND L. WELCH	
Effects of esthetic surroundings: I. Initial effects of three esthetic conditions upon perceiving "energy" and "well-being" in faces	247
BY A. H. MASLOW AND N. L. MINTZ	
The variability of reaction time in the sensori-motor induction syndrome with special reference to the effect of colors	255
BY L. HALPERN AND S. KUGELMASS	
A tactual-kinesthetic curvature illusion	271
BY EUGENE S. EDINGTON	
Regressive behavior changes in the tuberculous patient	273
BY SOL CHAREN	
A comparison between a hospitalized and a normal group on the X-O Tests	291
BY A. W. JONES	
The effect of organization upon complex reaction time	301
BY IVAN FLORES	
A study of authoritarianism and psychopathology	315
BY MERVIN FREEDMAN, HAROLD WEBSTER, AND NEVITT SANFORD	
Is mental illness mental?	323
BY GEORGE WATSON	
Space localization under conditions of danger	335
BY SEYMOUR WAPNER, HEINZ WERNER, AND PETER E. COMALLI, JR.	
The interaction chronograph as an instrument for objective measurement of interaction patterns during interviews	347
BY JOSEPH D. MATARAZZO, GEORGE SASLOW, AND RUTH G. MATARAZZO	
The collection and use of retrospective data	369
BY PAUL BAKAN	
The construction of an industrial selection personality test	379
BY WALTER V. CLARKE	
A suggested measure of bivariate relationship based on the ellipse of concentration	395
BY DONCASTER G. HUMM	
The personality profiles of loan office managers	405
BY WALTER V. CLARKE	
Personality profile of self-made company presidents	413
BY WALTER V. CLARKE	
The synonym vocabulary test: Standardization and validation	419
BY WILLIAM M. LEPLEY AND MARTIN L. ZEIGLER	
Reinforcement value as related to decision time	427
BY EDWIN J. LOTSOFF	
Subliminal perception and subception	437
BY JOSEPH H. VOOR	
Effects of esthetic surroundings: II. Prolonged and repeated experience in a "beautiful" and an "ugly" room	459
BY NORBETT L. MINTZ	

THREAT AS A FACTOR IN RECALL IN A RETROACTIVE PARADIGM*

Department of Psychology, Adelphi College

NORMAN M. CHANSKY¹

A. INTRODUCTION

Decrement in recall of an earlier learned task when a second task is interpolated has been explained in terms of confusion in the recall process (11), competition of a later learning with an earlier learning (7), and in terms of unlearning original material during the learning of the interpolated task (8).

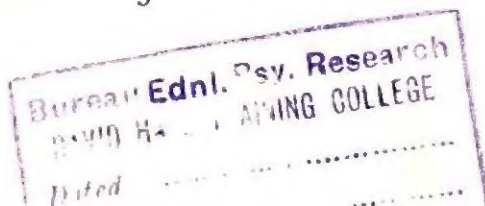
When two semantically similar tasks contiguous in time are presented, the earlier task is remembered less well. Gibson (2) points out that the decrement in recall occurs not only because of psychological generality, but more especially because of the need to make discriminations between two tasks. Osgood (9) adds that retroaction increases when each aspect of a task serves both as response to a preceding stimulus and as a stimulus to a succeeding item.

In addition to similarity and temporal contiguity, variables attending either the learning or the performance may cause a loss in recall. Of these "set" has been investigated for its possible effects in transfer experiments. Lester (6) found that information in the form of warning, directions, and hints tended to increase the retention of the initial learning. Postman and Postman (10) found retention of the initial learning increased when the set of the interpolated task was different from that of the original task. Similar results were obtained when an interpolated task was incompatible with the original learning. Irion (4), moreover, believing that forgetting is inherent in the classical design for the retroaction experiment, tested the effect of warm up before recall. The group with no rest before recall correctly anticipated more tasks than similar groups with a 24-hour rest. Of the groups with a 24-hour rest, the one in which the set was reinstated correctly anticipated more tasks than the one in which set was not reinstated.

To sum up briefly, although forgetting takes place in the retroaction ex-

*Received in the Editorial Office on August 16, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹The writer wishes to express his sincere appreciation to Dr. Irving Lorge of Columbia University Teachers College and to Dr. Marvin Iverson of Adelphi College for their helpful comments in preparing this manuscript.



periment, it can be decelerated or unaffected depending upon the information given when the set is instated.

An important variable yet to be intensively investigated is the "set" within the subject. His set includes such variables as attitudes, past experience, reaction to the experimenter, and general feeling tone at the time of the experiment. Precise measurements of these variables are not easily obtained.

Experiments in selective forgetting demonstrate that recall is better when there is ego-involvement. Shaw (12) reports that subjects who had been given bogus ratings on personality characteristics recalled more favorable than unfavorable ratings. Shaw and Spooner (13) report that ratings with which a subject agrees are remembered better than those with which he disagrees. Finally, Wallen (14) demonstrates that ratings on unidentified people were not recalled as well as ratings on acquaintances.

The impression that the last three studies give is that pleasant material is more easily recalled than unpleasant (unfavorable). Edwards (1), however, feels that more important than affectivity of an experience is what its meaning is to the subject. The subject interprets an experience in terms of his values, his frame of reference, and what he believes himself to be.

Recently a series of studies on the effect of the subjective meaning of a task upon recall have been reported. [Summaries are found in Lazarus et al. (5)]. Typical of these experiments are those of Glixman (3), who found that more completed than incompleted tasks were recalled under stress. This finding is contrary to expectations from the Zeigarnik Interrupted Tasks Situations (15). Zeller (16), in another type of task, demonstrated that induced failure at one task contiguous in time with another causes decrement in recall of the latter.

These experiments prompted the writer to design an experiment in which recall of an earlier learned task, in which decrement is inevitable, meant maintenance of status. The question to be answered was "What is the effect of threat on recall in a retroaction paradigm?"

B. PROBLEM

In the present study the effect of set before recall in a retroaction paradigm was investigated. The material to be recalled was a list of words with potential for generalization from the original list to the interpolated list. The sets induced before recall were threat and non-threat. The hypothesis to be tested was that recall under the two conditions is the same. The threatened group was told that through recall it was possible to determine intelligence and personality factors. Since all the subjects were students in educa-

TABLE 1
LISTS OF WORDS TO BE LEARNED

Card Pack I Original learning	Card Pack II Interpolated learning
cat	dog
man	woman
apple	peach
rabbit	squirrel
knife	dart
paint	draw
daughter	son
chicken	rooster

tional psychology courses (Columbia University Teachers College) it was assumed that their superficial acquaintance with psychological content would permit revelation of personality or intelligence factors to be threatening.

C. SUBJECTS

The subjects were volunteers from courses in Educational Psychology at Columbia University Teachers College. They were told that the experiment was one involving the learning of words. The placement of a volunteer in a group was decided upon before the experiment. The last 14 volunteers were to be classified Group *A*; the odd numbered signees were to be Group *B*; and the even numbered subjects were to be Group *C*. Forty-two students volunteered. Thirty-six students appeared for their appointments.

D. PROCEDURE

Before the experiment began each subject was told that the experimental procedure was a common one in psychology. He was then told:

Here is a pack of cards. Look at each word. When I tell you to turn, turn the card over and look at the next word. Turn that card over when you are so directed. This procedure will be followed until you go through the pack. Then, you will go through the pack again. This time, however, you will look at the word on the card and will anticipate the next word. You will go through the pack until you have learned all the words in the pack.

Learning was by the anticipation method. Each subject was allowed five seconds of study time. On the trial at which each succeeding word on the list was correctly anticipated, the list was said to have been learned.

When the original list was learned, each subject was told:

Here is a pencil and paper. Relax for the next few minutes by drawing whatever comes to mind.

After two minutes the subjects were told to stop.

At this point Groups *B* and *C* were given a second list to learn. They were told:

Here is another pack of cards. Look at each word. When I tell you to turn, turn the card over and look at the next card. Turn that one over when you are so directed. This procedure will be followed until you go through the pack. Then, you will go through the pack again. This time, however, you will look at the word on the card and anticipate the word on the next card. You will go through the pack until you have learned all the words in the pack.

This procedure is the same as the one before.

The criterion for learning again was the trial at which each word on the list was correctly anticipated.

Before recall, the two control groups (*A* and *B*) were told:

This is an experiment in learning. It is in no way influenced by intelligence or personality factors. It will demonstrate the effect of a rest period on the recall of material learned earlier.

Group *C*, however, was told:

This is a new projective test of intelligence. It was constructed to measure both one's native intellectual capacity as well as certain neurotic tendencies, such as insecurity, hostility, and aggression. It is possible to determine these by the number of correct responses made in the recall of the first list.²

After the set for recall was induced, all subjects were told:

Here is pack one. Go through the pack, looking at the word appearing on the card, saying what the next word in the pack is. When you have said the first word turn the card over immediately and do the same for the next word. Follow this procedure until you have gone through the pack.

The recall score was the number of correct anticipations of the first list, each response serving as a stimulus for each succeeding word on the list.

TABLE 2
SUMMARY OF DESIGN OF THE EFFECT OF THREAT ON RECALL

Group	Set prior to learning	Original learning	Two-minute rest activity	Interpolated learning	Set	Recall
A	Each subject told the experiment	List I	drawing	none	No threat	List I
B	Common in Psychology	List I	drawing	List II	No threat	List I
C		List I	drawing	List II	Threat	List I

²Each subject was individually interviewed. The entire experimental procedure was explained. This explanation served two purposes: 1. reduce threat and 2. satisfy curiosity regarding experimental procedure.

E. RESULTS AND DISCUSSION

Analysis of variance of the original learning of the three groups suggests that the groups come from the same parent population (Table 3). Although an analysis of variance of the performance of Groups *B* and *C* on the interpolated task suggest that Group *C* is significantly more variable than Group *B*, a comparison of the means of the two groups indicates that the differences between the groups may be due to random errors. Further evidence for the latter comes from the examination of the raw scores of Group *C*. One subject contributed to 78 per cent of the variance of the entire group.

TABLE 3
ANALYSIS OF VARIANCE TO DETERMINE WHETHER GROUPS ARE FROM THE SAME POPULATION

Group	Original list			Fma	Interpolated list			<i>F</i>	<i>t</i>
	<i>N</i>	Mean No. trials	Variance		Mean No. trials	Variance			
Group A	10	5.6	4.93		—				
Group B	12	4.6	8.77	2.07*	2.5	1.00			.57*
Group C	14	5.1	10.28		3.9	16.99	16.6**		

*Not significant at .05 level.

**Significant at .01 level.

Groups *B* and *C* learned the interpolated material in significantly fewer trials than they learned the original material. This may be due either to position generalization from the first list, e.g., domestic animal (first); human being (second); fruit (third) etc. or to learning to learn lists of words (Table 4).

TABLE 4
TRIALS REQUIRED TO LEARN ORIGINAL AND INTERPOLATED MATERIAL

Group	<i>N</i>	Mean difference in learning: Original minus interpolated	Standard error of the mean	<i>t</i>
Group B	12	1.8	.79	2.20*
Group C	14	2.5	1.05	2.3

*Significant at .025 level.

To answer the question "Does the interpolated learning interfere with the recall of an earlier learning?" it may be seen from Table 5 that the group not having the interpolated list correctly anticipated more words than the other groups. Since the only activity differentiating Groups *B* and *C* from Group *A* is the learning of the second list, it is concluded that the second learning caused the decrement in recall of the first. This writer chose to

explain this phenomenon in terms of probability. Since Group *A* was exposed to 7 stimuli and since Groups *B* and *C* were exposed to 14 stimuli, the number of combinations of responses for the former is fewer than for the latter. The likelihood that the latter would choose the correct response in recalling the original list is far less than for the former.

TABLE 5
A COMPARISON OF THE RECALL SCORES OF THE THREE GROUPS

Groups	N	Mean number of items correctly anticipated	<i>t</i>	Variance	Degrees of freedom	<i>F</i>
A	10	6.5				
B	12	5.0	3.6*			
A	10	6.5				
C	14	3.9	4.5**			
B	12	5.0		1.6	11	
C	14	3.9	1.83***	2.99	13	1.86****

*Significant at .01 level.

**Significant at .0001 level.

***Significant at .05 level.

****Not significant.

In other words, the fewer the responses available for a specific stimulus, the greater is the probability that the response elicited in recall is the correct one. This explanation is in line with the competition theory of reproductive (retroactive) inhibition (7).

The next question to be answered is "Does the set before recall influence the recall?"

One would expect no differences to occur between Groups *B* and *C*, since they both had identical learning experiences. The probability that Group *B* will respond correctly to a stimulus is exactly the same as that for Group *C*.

One can see from Table 5, however, that there are slight but significant differences between Groups *B* and *C* in the recall task. Since these groups are from the same parent population and since they had identical learning experiences (Table 3) the determining factor here is set before recall. To review, Group *B* was told that the experimenter was trying to determine the effect which the learning of a second task has on the recall of the first. It was emphasized that intelligence and personality factors were not being investigated. Group *C*, however, was told that the experimental situation was one through which it is possible to determine one's intelligence as well as certain personality characteristics.

An analysis of variance of the recall scores of Groups *B* and *C* reveals

that the threatened group was not significantly more variable than the non-threatened group. This finding suggests that threat before recall does not make the threatened group more variable, though it depresses the recall score.

It is the writer's belief that the decrement in recall of Group C resulted from the fact that the threatened group was not alerted in the learning task to what was to be expected of them in the recall task. Group B, on the other hand, expected a learning task according to the instructions given before learning. These instructions in no way conflicted with the instructions before recall. As a matter of fact, they were reassured that this was no intelligence or personality test.

The conflict in set before learning with that before recall seems to be the factor instrumental in depressing the recall scores of Group C.

F. SUMMARY AND CONCLUSIONS

An experiment on the effect of threat before recall was conducted. The three groups learned and recalled a list of words. Two groups learned a second list of words. There were no significant differences between the groups regarding learning of lists of words. When an interpolated list of words was learned, it impaired the recall of the original learning. When a threat was administered, moreover, it further impaired recall. The decrement in recall was explained in terms of the conflict aroused by the dissimilar cues in set before learning and before recall.

REFERENCES

1. EDWARDS, A. L. The retention of affective experiences: A criticism and re-statement of the problem. *Psychol. Rev.*, 1942, **49**, 43-53.
2. GIBSON, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196-229.
3. GLIXMAN, A. E. Recall of completed and incompleated activities under varying degrees of stress. *J. Exper. Psychol.*, 1949, **39**, 281-295.
4. IRION, A. L. The relation of set to retention. *Psychol. Rev.*, 1948, **55**, 336-341.
5. LAZARUS, R. S., DEESE, J., & OSLER, S. The effect of psychological stress upon performance. *Psychol. Bull.*, 1952, **49**, 293-317.
6. LESTER, O. P. Mental set in relation to retroactive inhibition. *J. Exper. Psychol.*, 1932, **15**, 681-699.
7. MCGEOCH, J. A. *The Psychology of Human Learning*. New York: Longmans, Green, 1942.
8. MELTON, A. W., & IRWIN, J. McQ. The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *Amer. J. Psychol.*, 1940, **53**, 173-203.
9. OSGOOD, C. E. The similarity paradox in human learning: a resolution. *Psychol. Rev.*, 1949, **56**, 132-143.

10. POSTMAN, L., & POSTMAN, D. L. Change in set as a determinant of retroactive inhibition. *Amer. J. Psychol.*, 1948, **61**, 236-242.
11. ROBINSON, E. S. The "similarity" factor in retroaction. *Amer. J. Psychol.*, 1927, **39**, 297-312.
12. SHAW, F. J. Two determinants of selective forgetting. *J. Abn. & Soc. Psychol.*, 1944, **39**, 434-445.
13. SHAW, F. J., & SPOONER, A. Selective-forgetting when the subject is not "ego-involved." *J. Exper. Psychol.*, 1945, **35**, 242-247.
14. WALLEN, R. Ego-involvement as a determinant of selective forgetting. *J. Abn. & Soc. Psychol.*, 1942, **37**, 20-39.
15. ZEIGARNIK, B. Das Behalten erledigter und unerledigter Handlungen. *Psychol. Forsch.*, 1927, **9**, 1-85.
16. ZELLER, A. F. An experimental analogue of repression: III. The effect of induced failure and success on memory measured by recall. *J. Exper. Psychol.*, 1951, **42**, 32-38.

Department of Psychology
Adelphi College
Garden City, New York

LYSERGIC ACID DIETHYLAMIDE (LSD-25): XVIII. EFFECTS OF LSD-25 AND SIX RELATED DRUGS UPON HANDWRITING*

*Departments of Medicine, Neurology, and Psychiatry, Mount Sinai Hospital,
New York*

M. W. HIRSCH, M. E. JARVIK, AND H. A. ABRAMSON^{1,2}

A. INTRODUCTION

Weyl (8) demonstrated that the handwriting of subjects about three to four hours after they received LSD-25 appeared large, illegible, and uncontrolled in contrast to their handwriting in the non-drug state. Incidental observations of handwriting and figure and Bender-Gestalt drawings of subjects tested in this laboratory indicate similar changes after LSD-25 ingestion. Results on a pursuit-rotor test and a steadiness test demonstrated that the drug decreased motor control (1).

Subjects under LSD-25 reported on questionnaires that they were unsteady, uncoordinated, and their movements lacked control (2, 4). When six drugs in addition to LSD-25 were studied, the results demonstrated that whereas little effect was produced by tap water placebos, a moderate effect on these functions was induced by drugs other than LSD-25, with a remarkably increased effect from lysergic acid diethylamide (4).

The purpose of this paper is to measure differences in the handwriting of subjects under the influence of a tap water placebo and seven related drugs: LSD-25 (lysergic acid diethylamide), LAE-32 (lysergic acid monoethylamide), BOL-148 (d-1-brom lysergic acid diethylamide), Ergonivine (ergometrine), ethyl alcohol, Pervitin or Methedrine (methamphetamine hydrochloride), and scopolamine (hyoscine). The relationship among these drugs is explained in another paper (4). The handwriting variables studied were suggested by the works of Castelnuovo-Tedesco (3), Roman (6), Saudek (7), and Wolfson (9).

*Received in the Editorial Office on August 17, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹This investigation was aided in part by a grant from the Josiah Macy, Jr. Foundation, New York City, New York.

²We are indebted to Sandoz Pharmaceuticals, Incorporated for supplies of LSD-25 and other compounds.

B. METHOD

1. Subjects

There were three male and two female paid adult volunteers, considered non-psychotic on the basis of a psychiatric interview and a Rorschach test. The median age was 26, ranging from 24 to 32 years. Body weight of the group ranged from 130 to 180 pounds, with the median at 150. There were three graduate students, a photographer, and an advertising trainee in the group.

2. Tests

On each experimental day the subjects copied verbatim a poem (5). They wrote the poem on a standard $8\frac{1}{2} \times 11$ inch sheet of paper. At the top of the page there were mimeographed blanks to be filled in with the subject's name, the date, the test, and the time.

3. Procedure

a. Drug administration. In administering the drugs, a double-blind procedure was followed wherein neither the subjects nor those administering the drugs knew which substance the subjects received on a given day.

In addition to the seven drugs studied, two tap water placebos were given. The five subjects were tested together during each experiment. On the first day all subjects received a placebo consisting of only 200 cc of tap water. During the remaining eight experiments, which were held at intervals of from two to six days, the drugs and a second placebo were each given once. No two subjects received the same drug in any one day. Table 1 indicates

TABLE 1
DRUGS RECEIVED ON EACH EXPERIMENTAL DAY BY EACH SUBJECT

Subject	Drugs								
	Placebo No. 1	LAE-32	Scopolamine	Placebo No. 2	Methamphetamine	LSD-25	BOL-148	Alcohol	Ergonovine
A	1	2	3	4	5	6	7	8	9
B	1	5	6	7	8	9	2	3	4
C	1	3	4	5	6	7	8	9	2
D	1	8	9	2	3	4	5	6	7
E	1	7	8	9	2	3	4	5	6

Note: Numbers 1 through 9 indicate the day on which the subjects received each drug.

the substance received by each subject on the given experimental day. The doses of the drugs were as follows: BOL-148 and LAE-32, .005 milligrams per kilogram body weight on the second experimental day and .007 milligrams per kilogram body weight on other days; methamphetamine, 7.5 milligrams per total body weight; LSD-25, .001 milligrams per kilogram body weight; alcohol, 45,000 milligrams per total body weight; Ergonovine, .65 milligrams per total body weight; and scopolamine, 1.3 milligrams per total body weight. Each drug was diluted in 200 cc of tap water. None of the drugs had any color and only alcohol had a taste and an odor. Subjects could not detect what substance they received.

The substances were administered orally at 9:30 A.M., except for the first placebo, which was given at 10:00 A.M. The subjects had eaten no food since the previous evening. One-half hour after the ingestion of the drug or placebo they ate a light breakfast. No smoking was allowed and no stimulants, such as coffee or tea, were permitted during the experimental period.

b. Test administration. The following instructions were given to the subjects: "Take a blank form and after *test*, write *handwriting*. Fill in the correct time. Now, copy this paragraph exactly as you see it here."

The test was given about $4\frac{1}{2}$ hours after the ingestion of the substances, except for the first day when it was given about 15 minutes after the placebo. The subjects took a variety of other tests before and after the handwriting test.

c. Analysis of data. Certain aspects of the subjects' handwriting in copying the poem were analyzed. These were: (a) total area used to copy the poem, (b) average height of capital letters, (c) number of errors in spelling, punctuation, and copying detail, (d) number of erasures, deletions, corrections, and cases of "mending," (e) slant of writing, (f) slant of first and of last line, (g) carefulness in dotting i's and crossing t's, (h) word spacing, (i) letter size, and (j) pencil pressure. The first four aspects were measured, whereas the others were estimated. The total area was that enclosed by lines drawn parallel to the sides of the sheet along the outermost extension of letters on each side. Each sample was analyzed individually and group trends were noted whenever they existed.

C. RESULTS

Tables 2 through 6 summarize the results of the four relatively objective measures made for each handwriting sample. The data for each subject appear on a separate table.

The total area used by Subject *A* ranged from 104.4 square cm to 121.0

square cm after receiving the two placebos, Ergonovine, and methamphetamine, as shown in Table 2. It ranked from 127.9 square cm to 179.3 square cm after he received the other drugs. When under the influence of LAE-32 and scopolamine the subject used the greatest space. There was

TABLE 2
QUANTITATIVE ANALYSIS OF HANDWRITING SAMPLES OF SUBJECT *A*: (A POEM WAS COPIED DURING 9 EXPERIMENTS)

Item	Substance received								
	Water placebos		Drugs						
	No. 1	No. 2	Ergonovine	Alcohol	Scopolamine	BOL-148	Methamphetamine	LAE-32	LSD-25
a. Total area, in square centimeters	104.4	121.0	118.8	149.5	168.6	147.0	127.4	179.3	154.7
b. Average height of capital letters, in centimeters	.6	.6	.6	.6	.7	.8	.6	.7	.7
c. Number of errors in spelling, punctuation, and copying	2	2	2	2	5	2	3	2	2
d. Number of erasures, deletions, corrections, and cases of "mending"	0	1	0	0	7	1	3	1	1

little variability in the average height of the capital letters. The greatest number of errors in spelling and punctuation and the greatest number of erasures, deletions, corrections, and cases of "mending," occurred after scopolamine; the second greatest number occurred after methamphetamine. The results after the other substances were fairly comparable. This subject wrote most irregularly after LSD-25 and scopolamine; his pencil pressure was greatest after receiving methamphetamine, LAE-32, and scopolamine. The size of the writing tended to correspond with total area. There were no other outstanding differences related to specific drugs.

Table 3 summarizes the analysis of Subject *B*'s handwriting. The total

area she used was generally greater than that used by other subjects. Under alcohol she used the smallest area and under the second placebo she used the greatest. Data from the first placebo experiment are missing. The average height of the capital letters was smallest after the placebo and large-

TABLE 3
QUANTITATIVE ANALYSIS OF HANDWRITING SAMPLES OF SUBJECT B: (A POEM WAS
COPIED DURING 9 EXPERIMENTS)

Item	Substance received								
	Water placebos		Drugs						
	No. 1	No. 2	Ergonovine	Alcohol	Scopolamine	BOL-148	Methamphet- amine	LAE-32	LSD-25
a. Total area, in square centimeters	*	391.0	301.0	252.4	324.7	364.5	343.1	359.8	343.1
b. Average height of capital letters, in centimeters		.9	1.2	1.1	1.4	1.3	1.0	1.4	1.0
c. Number of errors in spelling, punctuation, and copying		2	4	3	3	2	2	3	2
d. Number of erasures, deletions, corrections, and cases of "mending"		1	0	0	0	0	1	0	0

*Data missing.

est after scopolamine and LAE-32. The greatest number of errors was made after the subject received Ergonovine; the number of erasures and corrections never exceeded one. Subjective analysis of her handwriting size showed that it was small when she received the placebo and large when she received BOL-148, Ergonovine, and LSD-25. Pencil pressure was greatest after alcohol and LSD-25.

Subject C demonstrated considerable variability in the area of her writing after she received different substances. The total area ranged from 111.3 square centimeters after LSD-25 to 264.6 square centimeters after the second placebo. The area occupied after the first placebo fell approximately in

the middle of the range. The average height of the capital letters was least after LSD-25 and greatest after the second placebo. The greatest number of errors was made after alcohol; the greatest number of erasures, after scopolamine. Pencil pressure was heaviest after the second placebo and

TABLE 4
QUANTITATIVE ANALYSIS OF HANDWRITING SAMPLES OF SUBJECT C: (A POEM WAS
COPIED DURING 9 EXPERIMENTS)

Item	Substance received								
	Water placebos		Drugs						
	No. 1	No. 2	Ergonovine	Alcohol	Scopolamine	BOL-148	Methamphet- amine	LAE-32	LSD-25
a. Total area, in square centimeters	154.1	264.6	164.0	125.0	160.7	127.0	175.8	155.2	111.3
b. Average height of capital letters, in centimeters	.8	1.1	.8	.8	.8	.7	.7	.7	.6
c. Number of errors in spelling, punctuation, and copying	2	0	2	5	2	3	2	3	1
d. Number of erasures, deletions, corrections, and cases of "mending"	0	0	0	2	3	0	0	0	0

after scopolamine. The writing was most careless and the spacing between words was closest when the subject was under the influence of alcohol.

The total area after different drugs was fairly consistent for Subject *D* in comparison with other subjects. It ranged from 152.9 square centimeters after the first placebo, to 220.5 square centimeters after LAE-32 and BOL-148, as shown in Table 5. Average capital letter size also varied little, but correlated with total area. There was one error made after scopolamine and one after LAE-32; otherwise there were no errors. There were 6 erasures after the subject received scopolamine and 3 after he received methamphetamine and LSD-25. He wrote largest after alcohol, BOL-148 and LAE-32. The writing always slanted toward the right, except after

he received BOL-148 and LSD-25, when the slant was irregular. Pencil pressure, although fairly consistent, was slightly lighter after the ingestion of Ergonovine and alcohol.

Subject *E* exhibited the greatest variability in total area used after the

TABLE 5
QUANTITATIVE ANALYSIS OF HANDWRITING SAMPLES OF SUBJECT D: (A POEM WAS COPIED DURING 9 EXPERIMENTS)

Item	Substance received								
	Water placebos		Drugs						
	No. 1	No. 2	Ergonovine	Alcohol	Scopolamine	BOL-148	Methamphetamine	LAE-32	LSD-25
a. Total area, in square centimeters	152.9	177.5	*	192.9	165.5	220.5	166.3	220.5	203.9
b. Average height of capital letters, in centimeters	.7	.8	.8	.9	.9	.9	.8	1.0	.8
c. Number of errors in spelling, punctuation, and copying	0	0	0	0	1	0	0	1	0
d. Number of erasures, deletions, corrections, and cases of "mending"	0	0	0	1	6	0	3	0	3

*Data missing.

different substances and showed the most pronounced handwriting change after LSD-25, when he used 496.8 square centimeters area; he had to use the reverse side of the page to complete the poem. He wrote smallest under the influence of LAE-32 where the total area used was 184.9 square centimeters. Capital letter size varied from .6 centimeters after the first placebo to 2.0 centimeters after LSD-25. The greatest number of errors was made after LSD-25, with a large number also occurring after the two placebos and scopolamine. Many erasures were made only after scopolamine. After LSD-25, the handwriting was most irregular and uncontrolled. Spacing between words was inconsistent; the subject was careless in dotting *i*'s and

crossing *t*'s, and wrote with the heaviest pencil pressure. After the placebos and LAE-32, pressure was lightest.

The handwriting protocols of two subjects after they received the second placebo and after they received LSD-25 appear in Figures 1 and 2. They

TABLE 6
QUANTITATIVE ANALYSIS OF HANDWRITING SAMPLES OF SUBJECT E: (A POEM WAS COPIED DURING 9 EXPERIMENTS)

Item	Substance received								
	Water placebos		Drugs						
	No. 1	No. 2	Ergonovine	Alcohol	Scopolamine	BOL-148	Methamphetamine	LAE-32	LSD-25
a. Total area, in square centimeters	215.3	235.6	215.8	253.5	226.0	300.1	228.8	184.9	496.8
b. Average height of capital letters, in centimeters	.6	1.2	.9	1.0	.9	1.1	1.1	.8	2.0
c. Number of errors in spelling, punctuation, and copying	4	5	2	1	4	1	1	2	7
d. Number of erasures, deletions, corrections, and cases of "mending"	1	0	0	0	6	0	1	1	1

demonstrate more vividly the kind of changes which can occur. These two subjects showed opposite tendencies when under the influence of LSD-25.

D. DISCUSSION

Despite the differences among subjects there are some generalizations which can be made about the handwriting changes occurring after the subjects receive certain drugs.

It is interesting that four subjects for whom data are available used a greater amount of space to copy the poem after the second placebo than after the first. This may be attributed to practice and decreased attention to the

task. The line slant of the first line and of the last line was the same in 8 out of 10 placebo experiments.

Scopolamine tends to lead to a very large number of erasures, deletions,

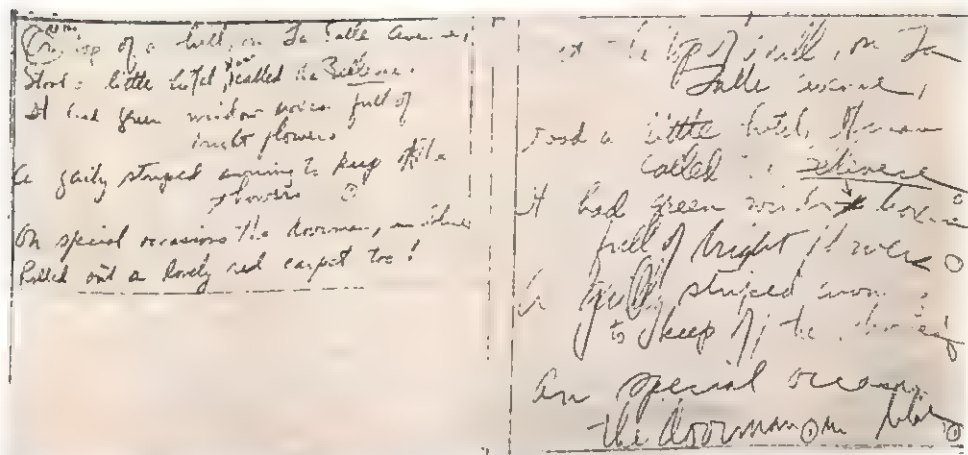


FIGURE 1
HANDWRITING PROTOCOL OF SUBJECT E, OBTAINED AFTER PLACEBO INGESTION (LEFT)
AND LSD-25 INGESTION (RIGHT)

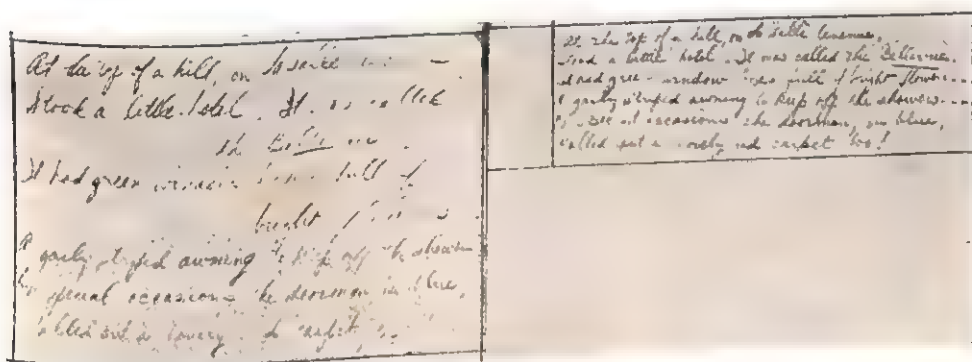


FIGURE 2
HANDWRITING PROTOCOL OF SUBJECT C, OBTAINED AFTER PLACEBO INGESTION (LEFT)
AND LSD-25 INGESTION (RIGHT)

and corrections. This drug also produces extreme feelings of drowsiness in the same subjects (4), and drowsiness may interfere with the subjects' alertness and ability to perform efficiently. In four subjects there was a great deal of irregularity in form, letter size, word spacing, and slant of writing. The relative size of the handwriting and the area used after the subjects received

this drug varied; for some it was smaller than after most other drugs; for others, larger.

The handwriting of most subjects was relatively large under BOL-148 and relatively small under Ergonovine. Under these drugs and alcohol the subjects were able to maintain a consistent line slant. Under Ergonovine there were no erasures, etc., by any subject and under BOL-148 only one subject had an erasure. There was greater variability among the other drugs along these parameters.

The area used by three subjects was larger after they received LSD-25, BOL-148, and alcohol than after the placebos. Two subjects used the largest area under the second placebo and used little space under alcohol and LSD-25, showing opposite effects of these drugs. The relative positions of the placebo and other drugs was random and inconsistent.

Irregularity in letter size, word spacing, slant and lack of control in writing seem to characterize the writing of four subjects while under the influence of LSD-25. These changes did not seem to be present in the writing of one subject who wrote smallest under this drug. The slant of the first and of the last line differed for each of four subjects under LSD-25.

Subject *C* made more errors after alcohol than after any other drug, and more errors than other subjects made after alcohol. She also gave more positive responses to the questionnaire (4) than other subjects, suggesting that alcohol had a comparatively great effect upon her. Two subjects displayed irregularity in writing after alcohol.

There was some correlation between the handwriting changes and the number of positive questionnaire responses of Subject *C* after alcohol. For the group as a whole, the questionnaire showed that subjects gave relatively few responses following the water placebos and gave the greatest number of responses after LSD-25. No such clear-cut differentiation among substances was found in these handwriting samples.

Calling this test a *handwriting* test might motivate subjects to compensate for possible changes in their writing. Comparison of these handwriting samples with samples where memory and other factors were of major interest showed that, regardless of the test, the handwriting on a given day was fairly consistent.

E. SUMMARY AND CONCLUSIONS

Five non-psychotic, adult subjects received average effective doses of LSD-25, LAE-32, BOL-148, Ergonovine, alcohol, scopolamine, and methamphetamine, administered orally in 200 cc of tap water. They also received two tap water placebos. Each substance was given on a separate day; no

two subjects received the same drug the same day. About $4\frac{1}{2}$ hours after receiving each substance the subjects copied a poem. Their handwriting was analyzed for total area, the average height of the capital letters, the number of errors in spelling, punctuation, and copying detail, and the number of erasures, deletions, corrections, and cases of "mending." Several subjective measures were also made; namely, letter size, line slant, pencil pressure, etc. The following tentative conclusions are made:

1. Subjects demonstrated considerable, individual variability in their handwriting, after receiving the different substances.

2. Under the influence of LSD-25, alcohol, and BOL-148, three subjects wrote much larger than after the placebos, and two wrote smaller. Under Ergonovine subjects generally wrote small.

3. The greatest number of errors and the greatest number of erasures, deletions, corrections, etc., tended to occur under scopolamine. This substance produces feelings of extreme drowsiness which may be a causal factor in the number of errors and corrections. No subject made any erasures, etc., after receiving Ergonovine, and only one subject made an erasure after BOL-148.

4. The greatest irregularity and carelessness in writing occurred after alcohol, scopolamine, and LSD-25, with the most noticeable effects related to LSD-25 ingestion. The other drugs had no widespread marked effects.

5. There was little relationship between the handwriting of subjects and subjective changes under the influence of the two placebos and the seven drugs. Subjective changes were measured by their responses to a questionnaire.

REFERENCES

1. ABRAMSON, H. A., JARVIK, M. E., & HIRSCH, M. W. Lysergic acid diethylamide (LSD-25): VII. Effect upon two measures of motor performance. *J. of Psych.*, 1955, **39**, 455-464.
2. ABRAMSON, H. A., JARVIK, M. E., KAUFMAN, M. R., KORNETSKY, C., LEVINE, A., & WAGNER, M. Lysergic acid diethylamide (LSD-25): I. Physiological and perceptual responses. *J. of Psych.*, 1955, **39**, 3-60.
3. CASTELNUOVO-TEDESCO, P. A study of the relationship between handwriting and personality variables. *Genet. Psychol. Monog.*, 1948, **37**, 167-220.
4. JARVIK, M. E., ABRAMSON, H. A., & HIRSCH, M. W. Lysergic acid diethylamide (LSD-25): Comparative Study of Subjective Changes Under Seven Drugs (in preparation).
5. PARKIN, R. *The Red Carpet*. New York: Macmillan, 1948. Pp. 40.
6. ROMAN, K. *Handwriting, a Key to Personality*. New York: Pantheon, 1952. Pp. 382.
7. SAUDEK, R. *Experiments With Handwriting*. London: Allen & Unwin, 1928. Pp. 394.

8-322

8. WEYL, B. Versuch einer psychopathologischen Analyse der LSD-Wirkung. München, 1951. Pp. 61. (Albert-Ludwig U., M.D. thesis.)
9. WOLFSON, R. A Study in Handwriting Analysis. New York: Columbia Univ., 1949. Pp. 95. (Columbia U., Ph.D. thesis.)

133 East 58th Street

New York City 22

TWO KINDS OF CONFORMITY: A STUDY OF THE RIESMAN TYPOLOGY APPLIED TO STANDARDS OF PARENTAL DISCIPLINE*

*Center for Advanced Study in the Behavioral Sciences, Stanford, and
Department of Psychology, Boston University*

ARTHUR J. BRODBECK, PHILIP NOGEE, AND ALBERT DI MASCIO¹

A. INTRODUCTION AND PROBLEM

Social psychologists are continually faced with the problem of determining the norms that govern the way people interact. There is at least one ambiguity in the use of the word "norm" however, which conceivably masks some very interesting and important problems regarding conformity behavior. Sometimes, the word is used to designate what people use as an ideal, in terms of which they judge the "rightness" or "wrongness" of their social practices, whereas at other times the word is used to designate what people judge to be the real or actual practices other people around them, in similar social rôles to theirs, employ in their day to day social interactions. The first type of norm partakes more of the character of conscience, whether that be thought of as a superego or ego-ideal mechanism, brought into being by early socialization experiences and identifications established in primary groups. The second type of norm is more the result of a quasi-realistic "statistical" appraisal of the manifest behavior around one. Both types of norms may presumably act as sanctions to reinforce and strengthen habits of interpersonal participation, but they may often conflict. Conformity to an internal norm may jeopardize one's position with contemporary groups, and conformity with the behavior one sees about one may result in feelings of personal guilt and uneasiness. We have only an imperfect knowledge of how conflicts like these are instigated and how they are resolved and with what immediate and long-term individual and social consequences. Redfield (5) and Lewis (3), both observing the same social group, obtain conflicting anthropological evidence regarding the culture in which their field

*Received in the Editorial Office on September 1, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹This is one of a series of studies on conformity behavior by the senior authors. The suggestions of David P. Ausubel, Harold D. Lasswell, and Paul Lazarsfeld regarding the interpretation of the results have been very useful to us.

work was done, presumably because one report was based on ideal norms, whereas the other was based on observations of what people actually did. Recently, Bauer (1) in a paper on the Soviet middle-class elite, suggests that strong personality conflicts can be induced by too wide a felt divergence between actual and idealized norms, a state of affairs which can act to weaken social organization.

This same sort of problem has been brought forward very strongly in a series of writings by David Riesman and his associates (6). Riesman has attempted to discuss and document the hypothesis that social behavior in America has increasingly become more a process of "doing what others do" and less of thinking for oneself and following the dictates of one's own conscience. It is as if there were a premium on "flexibility," "being fashionable," and "moving with the crowd." In this process, it appears that internal standards have presumably lost some of the power to influence social conduct which they once possessed; a "social radar set" has been substituted for the "voice of conscience." In Riesman's terms, people have become more "other-directed" than "inner-directed."

Because Riesman's hypotheses are so global in nature, and stated in such an intuitive way, they are difficult to test, for there always remains, after the most careful reading, a certain ambiguity about meaning, and hence, of empirical indicators. It is not entirely clear why the shift from "inner-direction" has taken place, and thus, which professions and groups have been most affected. And while a great deal of materials has been amassed to indicate the shift that Riesman has in mind, the detailed differences in conformity and social interactions characterizing the "other-directed" and "inner-directed" are not explicitly etched, but are left implied and buried among a wealth of connotations for others to abstract. For these reasons, the relationship of Riesman's writings to empirical research is more in the nature of opening up an interesting new area for research study, during which the major hypotheses may be sharpened and refined by a sort of back and forth movement of theory and data-collection, rather than proved or refuted as they now stand. A detailed analysis of these concepts in *The Lonely Crowd* has been made elsewhere by one of the authors (2).

It is with such an aim in mind that we have attempted to gather some systematic data regarding the relative influence of "inner" and "other" norms on social behavior, for which we have attempted to supply operational indicators covering Riesman's terms. We have selected parental behavior to study, since child training figures so decisively in Riesman's thinking, both because it produces the character types, as well as reflects them. The normal

deliberations parents experience in socializing their children should, if Riesman is correct, be resolved with less conflict by conforming to what other mothers are felt to do in similar situations than by conforming to their own independent thought and moral standards. It should be noted that this last hypothesis is a systematic rather than an historical version of Riesman's idea. People might be found to have more inner-directed than other-directed tendencies at present, and yet they may still be more other-directed than they once were. Our study does not attempt to clarify the historical aspect of other-direction, but only the contemporary strength of these two tendencies.

A second and subsidiary question regarding social norms is also investigated in this study. General norms governing social relations, whether influenced by inner or external factors, are likely to vary somewhat according to the specific social situations and human behavior with which a person must cope. For instance, a majority of people living in Great Britain may feel that in general it is quite wrong to ever "make a scene." It is nevertheless quite conceivable that there may be special situations within which a display of temperament and anger would be generally thought quite permissible and even advisable. So, too, the norms governing parental discipline of children might conceivably vary in the same way. The design of our study allowed us to attempt to get answers to two such questions. Is one form of discipline more *generally* given preferential status among parents than all others? If so, what is the nature of the special conditions under which such preferred discipline techniques begin to lose preferential status in parent-child relationships?

B. RESEARCH DESIGN

Thirty mothers having at least one child between the ages of $3\frac{1}{2}$ to $7\frac{1}{2}$ comprised the subjects of our study. All of our informants lived in the same block of a veteran's housing project in one of the Boston suburbs. One of the requirements of the project was that it would accept for residence no family which had an income of more than \$4,000.00 a year, thus making for a high uniformity of the socio-economic level of our subjects. The average vocation is that of a skilled or semi-skilled worker. The families' length of residence is from $2\frac{1}{2}$ to 3 years, and the group is not very socially mobile, remaining in general in the same status. Very few of the mothers in the project work in order to bring in additional income. Thus, our population was one which probably could be described as "upper-lower" or "lower-middle" in social status.

Each mother was visited in her home by one of the authors, and was asked to respond to a self-administered multiple-choice questionnaire in the presence of the interviewer. The mothers were told that this was part of a survey being conducted by Boston University. The questionnaire was divided into three parts. Each part consisted of exactly the same 39 items which dealt with a variety of situations which mothers might be expected normally to encounter in the course of the socialization of their children. During the first administration, the mothers were instructed to answer the items in terms of what they *actually did* in these situations. During the second administration, the instructions were to respond in terms of what she thought she *should do*; and, then, finally, what she thought her *neighbors usually did*. The second administration was used to determine the internal norm; the third to determine the external (group) norm. The respondents were asked to give their actual behavior first, so as to minimize the influence of the norms, external or internal, on actual behavior in the reporting process.

The mothers indicated their response in each case by encircling any number from 1 to 6 for each item. The "meaning" of the numbers was explained to each mother before beginning each administration, and the code was available for them to look at during any administration. The number code was as follows:

1. Paddle or slap the child.
2. Scold the child.
3. Pay no attention to the behavior of the child.
4. Make the child stay alone for a while.
5. Talk it over with the child.
6. Tell the child that God will punish him.

These alternatives were selected because, although they are crude designations of what are obviously more complex kinds of parental behavior, they nevertheless are abstract enough to be applicable to all the situations on the inventory. Although the sixth alternative was used by mothers in a small pilot study, none of these particular mothers checked it on any administration. For this reason, it was dropped from the analysis. It should be emphasized that there was no attempt to include any particular type of special situations in the questionnaire, but instead as wide a variety of situations as possible was used. An attempt was made to minimize any "other-directed" tendencies toward the interviewer by emphasizing that we did not know which methods would give best results and were conducting the study for that reason.

C. RESULTS

Table 1 shows the frequency rank with which the various methods of disciplining children were selected. A method of discipline which received a rank of 1.0 would indicate that all the mothers *preferred* this method of discipline to all others; a method of discipline which received a rank of 5.0 would indicate that all the mothers disapproved of using this method more than all the others.

TABLE 1
MEAN RANK OF FIVE DISCIPLINE TECHNIQUES*

	1	2	3	4	5
Actual norm	3.2	2.8	3.0	4.5	1.4
Internal norm	3.1	2.9	3.0	4.5	1.4
External norm	2.4	2.0	3.0	4.5	2.6

*Mean rank was obtained by recording for each mother the number of times she circled each category, and then ranking the categories according to their frequency of use. The ranked categories for each subject were then summated and the average rank of use for each category was obtained.

It is quite evident that a strong and consistent *discipline taboo* seems to be operating among these mothers. Isolating techniques seem to be quite generally frowned upon, the disapproval far outranking that felt for any other method of discipline. There is no discrepancy between the rank order of internal and external norms in this regard, but instead perfect agreement among the two sets of norms regarding it. Furthermore, these mothers report that they avoid employing this method more than any other when they discipline their children, and, then, *exactly as infrequently* as both the external and internal norm require. It would appear that, among mothers approaching the lower end of a socioeconomic distribution, the pressure is quite powerful to refrain from using more indirect discipline techniques of "isolation" and implied "loss of love." Instead, the philosophy seems to be: discipline directly by reason or open expressions of anger or don't discipline at all and simply let the problem episode go unnoticed.

From the point of view of the Riesman hypothesis, the perfect agreement between internal and external norm regarding this discipline taboo could be interpreted in two possible ways. The agreement may itself be the result of a process whereby the mothers are unwilling to view their own internal norms as diverging or differing from the norms reflected in their neighbor's behavior. Their internal standards regarding discipline sanctions and taboos are, in other words, *suppressed* when they conflict with what everyone else is doing, and the external norms are substituted for them and introjected.

On the other hand, the result might reflect something quite different and consistent, too, with the Riesman hypothesis. It may be that norms of conduct have not changed much, but only the tendency to get "relief" when one's internal standards are confirmed by other people's behavior. The differences, in other words, are primarily in the value placed upon the *source* of reinforcement—self-reinforcement vs. group-reinforcement—rather than in the values being reinforced themselves; it is possible to follow the internal norms without anxiety only when it can be seen that other people follow them too, whereas once the reinforcement came entirely within. Our data do not permit us to deny or confirm either of these hypotheses, and we are uncertain as to how to relate them to the details of Riesman's writings.

A glance at the comparisons regarding internal and external norms for the other discipline alternatives makes either hypothesis quite untenable for the remainder of the data, however. There are quite marked discrepancies between the internal and external norm for the other four discipline techniques. In each case, parents report that their own behavior has been more heavily influenced by the internal norm than it has by the external norm. Furthermore, the internal norm is one that gives strong and clear-cut sanction to the use of "discussion" in socializing children. Although the ranking indicates that individual differences must be taken into account, the general thing to do, according to the internal norm of these mothers, is to resort to reason or logic in disciplining children in preference to more emotional forms of discipline. The external norm, however, offers no clear-cut sanctions. Other parents are perceived as resorting to discussion and physical punishment about equally, and to resort to scolding a little more than either. The rank values are very similar for all three discipline techniques, at least more similar than they are in terms of the image of the internal norm these mothers have reported. Two conclusions seem apparent. Where internal discipline sanctions differ from the external counterparts, the mothers in this study report themselves as decidedly inner-directed. Secondly, the sanctions provided by the external norm, apart from the fact that they generally depart from the sanctions emanating from the internal norm and are not as important in determining what parents do, seem to have no *general* clear-cut and compelling features when compared with the sanctions provided by the internal norm. It is as if these mothers perceived no general or well-organized philosophy of child care reflected in their neighbor's behavior. The reasons for this are, of course, still problematical. It may be that each parent only sees the other parent disciplining in special situations out of doors and extrapolates from these incorrectly to situations which occur in the pri-

vacy of the home. It may be that there is little identification with the "neighbors" which each parent reports about and, because of this, the perceived response of "other" parents toward their children takes on an out-group character. Our data do not permit answers to the "meaning" of the finding of strong inner-directedness for our informants, but does suggest an avenue of further research.

There is another way by which one can try to gauge the degree to which others act in accordance with inner or other established norms. Among the various socialization problems presented to each informant, it is possible simply to count the number of situations in which each parent reports her own behavior to be the same as that required by the internal norm, as well as the number of situations in which the report was the same as that required by external standards. When an item can be scored in both ways, it is scored both ways. For each subject, therefore, it is possible to obtain the relative strength of inner-directed and other-directed tendencies by this procedure. These data are summarized in Table 2 for each subject. In 24 of the 30 cases, the mother reports her behavior more often influenced by internal standards than it is by external standards. In five of the remaining six cases, the mothers are influenced equally in both directions. Only one mother (Informant 12) was slightly more other-directed than inner-directed. In terms of this kind of subject classification into the two categories, the evidence again indicates that these mothers are more strongly influenced by internal standards than they are by the wish to conform to standards other parents suggest and provide.

What are the reasons which cause mothers to "break with" either internal or external standards? It is possible to get some answer to this question by determining whether the deviation from either norm is consistently in the direction of more or less severe kinds of discipline than the norm dictates. These data are summarized in Table 3 subject by subject. In 14 of the 30 cases, the mothers break with the internal norm more often to resort to more severe methods of discipline. But an exactly similar number break with the internal norm for the opposite reason. The data seem to indicate that there is no general tendency among the mothers to consider their internal norms as either too severe or too lenient but that this factor, if it is present at all, varies considerably from mother to mother. This does not mean, of course, that personality factors which produce tendencies to be lenient or very strict do not operate to make individual mothers deviate from their internal norms of discipline. On the contrary, it means that idiosyncratic personality differences, not highly associated with class and status posi-

tion, are probably more important in inducing deviations from internal standards than is some kind of general social tendency affecting all mothers uniformly. The socialization experiences of mothers which produce tendencies in them to behave either more or less in accordance with their internal standards of discipline are themselves still an area of research which needs empirical study. Our findings merely suggest that these tendencies are relatively unassociated with class position—insofar as our population is representative of any general social class.

TABLE 2
THE RELATIVE FREQUENCY OF CONFORMANCE WITH INTERNAL AND EXTERNAL
NORMS BY SUBJECT

Subjects	Conformance with internal norms	Conformance with external norms
1	28	3
2	24	24
3	25	22
4	24	15
5	28	10
6	21	21
7	26	26
8	31	15
9	28	11
10	21	21
11	29	24
12	24	25
13	24	23
14	26	15
15	24	15
16	27	4
17	24	12
18	20	20
19	22	14
20	31	28
21	28	24
22	26	18
23	21	12
24	29	24
25	29	26
26	18	13
27	26	22
28	30	28
29	30	18
30	21	18

A comparison of the deviation from external standards, however, gives a much different kind of picture. In 22 of the 30 cases, the mothers deviate from the external standards more frequently in order to be more lenient. Despite the fact that the strongest normative feature provided by the image of external standards was the disapproval of isolating techniques,

these parents still find the external norm too frequently harsh to be guided by it.

TABLE 3
THE RELATIVE FREQUENCY OF DIRECTION OF SEVERITY

Subjects	Internal norm		External norm	
	Deviation toward more severe behavior	Deviation toward less severe behavior	Deviation toward more severe behavior	Deviation toward less severe behavior
1	9	2	19	17
2	8	7	12	3
3	8	6	5	12
4	8	7	6	18
5	6	5	14	15
6	9	9	8	10
7	7	6	6	7
8	3	5	5	19
9	5	6	18	10
10	10	8	8	10
11	4	6	9	6
12	8	7	9	5
13	11	4	8	8
14	5	8	6	18
15	9	6	4	20
16	10	2	0	35
17	6	9	9	18
18	5	14	8	11
19	8	9	12	13
20	3	5	1	10
21	5	6	8	7
22	2	11	10	11
23	9	9	8	19
24	7	3	11	4
25	4	6	4	9
26	7	14	1	25
27	7	6	6	11
28	3	6	0	11
29	8	1	1	20
30	7	11	4	17

Taken together, these results seem to imply that the informants were not only less influenced by external norms than they were by internal norms, but that they also share a common kind of critical attitude toward the external norm, which led them to a more uniform kind of deviation in lenient directions from it. The uniformity suggests, in other words, that the result is not due to idiosyncratic factors, as it seems to be for internal norms, but to a general kind of social tendency to criticize external norms within a certain common frame of reference—i.e., *other parents discipline too harshly*. Such a consistent critical frame of reference toward others is hardly compatible with "other-directedness."

In general, the data have shown that the internal norm is one that gives high sanction to forms of discipline involving reason rather than emotion and which strongly taboos isolation techniques in favor of a more open show of anger and displeasure. Except for the taboo on isolation techniques, the external norm diverges markedly from the internal norm. With the knowledge of the general image these mothers have of both internal and external norms now available to us, it is possible to see how these general norms vary according to the situations in which their operation was reported. For each situation it is possible to get the modal view of both external and internal norm, as well as the modal discipline behavior which these mothers reported using in that situation. These results are presented in Table 4.

TABLE 4

Item	Situation	Actual modal behavior	Modal internal behavior	Modal external behavior
1.	Wets pants at night	5	5	1
2.	Refuses to eat	3	3	2
3.	Refuses to go to sleep	5	5	2
4.	Gets clothes dirty	3	5	2
5.	Plays with parents' books	3	5	2
6.	Hits other children	5	5	1
7.	Afraid of strange things	5	5	5
8.	Gets into mischief while helping	5	5	2
9.	Hits himself	5	3	2
10.	Disobeys	1	1	1
11.	Afraid to be alone	5	5	5
12.	Breaks his toys	2	2	2
13.	Scratches furniture	1	1	1
14.	Bites nails	5	5	2
15.	Cries very easily	3	3	2
16.	Afraid of the dark	5	5	5
17.	Plays in street	1	1	1
18.	Gets hands and face dirty	2	3	3
19.	Has temper tantrums	1	1	1
20.	Writes on walls	2	1	1
21.	Picks nose	5	2	2
22.	Ignores what he is told to do	1	1	1
23.	Afraid of animals	5	5	5
24.	Hits brothers or sisters	2	2	1
25.	Sucks thumb	3	3	2
26.	Touches his sex organs	5	5	5
27.	Disobeys teachers	2	2	2
28.	Vomits	5	5	2
29.	Whines	2	2	2
30.	Doesn't allow children use of toys	5	5	2
31.	Questions constantly	5	5	2
32.	Afraid of lightning	5	5	5
33.	Wets pants during the day	1	5	5
34.	Doesn't want mother to leave	5	1	2
35.	Uses dirty words	5	5	2
36.	Refuses to play by himself	5	1	1
37.	Splashes water in the bathroom	2	3	3
38.	Spits at others	1	2	2
39.	Refuses to go to school	5	1	1
			5	2

The distribution of the items just happens to allow for three predominant classifications, none of which were pre-planned but arise clearly out of the empirical data. In two of the classifications, the modal response sanctioned by the internal and external norm agree either with respect to the necessity for punishment (alternatives of discipline 1, 2, or 4) or with respect to the necessity for non-punishment (alternatives of discipline 3 and 5). The situations comprising Group I, then, consist of norm agreement regarding punishment, and the situations comprising Group II consist of norm agreement regarding non-punishment. In both cases, the parents modally follow the convergent norms in reporting their actual behavior. Group III, however, is composed of situations in which the internal norm is morally seen as sanctioning non-punishment (and the modal actual response is non-punishing), but the external norm sanctions punishment. There are 13 situations in Group I, 8 situations in Group II, and 15 situations in Group III. This classification system allows us to examine the specific details of internal and external norm agreement and disagreement, as well as to get at the conditions under which the general sanction toward the use of reason in disciplining children is abandoned. The situations so classified into three groupings are listed, for the convenience of the reader, by description in Table 5. Only three situations do not fall into these three classifications.

TABLE 5

Group I Negativism and aggression	Group II Fear and anxiety	Group III Dependency, etc.
10. Disobeys	7. Afraid of strange things	1. Wets pants at night
12. Breaks his toys	11. Afraid to be alone	2. Refuses to eat
13. Scratches furniture	16. Afraid of the dark	3. Refuses to go to sleep
17. Plays in street	23. Afraid of animals	4. Gets clothes dirty
19. Has temper tantrums	26. Touches his sex organs	5. Plays with parents' books
20. Writes on walls	31. Questions constantly	6. Hits other children
22. Ignores what he is told to do	32. Afraid of lightning	8. Gets into mischief while helping
24. Hits brothers or sisters	36. Refuses to play by himself	9. Hits himself
27. Disobeys teachers		14. Bites nails
29. Whines		15. Cries very easily
33. Wets pants during day		25. Sucks thumb
37. Splashes water in the bathroom		28. Vomits
38. Spits at others		30. Doesn't allow children use of toys
		34. Doesn't want mother to leave
		39. Refuses to go to school

A glance at the situations falling under Group II reveals that they tend to be *fear or anxiety items*—i.e., situations in which the child shows fear of animals, the dark, lightning, being alone, etc. Despite some exceptions, the trend is quite obvious. In such situations, the mothers do not perceive other mothers acting differently from themselves and perceive instead a high degree of similarity between the internal and external norm. The philosophy seems to be that a child's fears or anxieties must be either handled with a large dose of rational discussion, or else not accentuated by special parental attention; in no event is the child to be physically punished, scolded, or isolated for the expression of fear and anxiety.

The situations falling under Group I also seem to have a dimension of similarity which runs through them. A large number refer to problems of negativism and aggression that parents face in the socialization of their children—i.e., writes on walls, hits siblings, disobeys mother, has temper tantrums, etc. Where the child challenges the parent's authority or attempts to hurt members of the immediate family, both the internal and external norm converge to sanction the parent's right to punish. At such times the philosophy seems to be that discussion and reason is useless and worthless; yet, the behavior cannot be ignored and the parent must resort to less preferred forms of discipline. This finding is quite important since it provides some clues, among these mothers at least, as to the conditions under which the discipline norm of discussion and reason breaks down. It is interesting that disobedience toward teachers is looked upon in the same way, and it would tend to suggest, as Stendler (8) has reported, that parents look upon school as an extension of the home, and orient their children in this way toward the school and teacher.

It is more difficult to find some common dimension among the situations falling under Group III, the situations in which external and internal norms are disparate. In a sense, this would seem to indicate that mothers, when they act in accordance with external norms, do so only in very specific kinds of situations whereas their divergence from external norms is a more diffuse tendency. To put it a little differently, these mothers are inner-directed in a wide variety of situations and perceive themselves acting like others only in quite special instances in which, furthermore, the external norms do not conflict with the internal norms.

However, this does not mean that there is no rhyme or reason to the situations falling under Group III, despite their greater lack of uniformity as compared with the situations falling in the other two classifications. The situations, on the contrary, seem to fall into three sub-groups. One sub-

group concerns *the child's unwillingness to do what is good for him*—i.e., refuses to eat, to sleep, hits himself, refuses to go to school, etc. The internal norms of these mothers sanction the use of reasoning for such behaviors or, perhaps, the alternative of letting him learn the “natural” consequences of refusing to do what is good for him. Neighbors are perceived as sanctioning punishment for such refusals, but the internal norm taboos such drastic measures. Instead, there seems to be a kind of faith that the child will “naturally” discover the consequences of his unwillingness and that will be punishment enough.

Another sub-group of situations seems to concern *the child's relationships with children outside the family*—i.e., hits other children and refuses to share toys. Our informants apparently felt that the child's relationships outside the family in the peer group are either something he must take responsibility for himself, or when more is required, should be dealt with by the permissive procedure of discussion. Neighbors, however, are perceived as too punitive in their handling of their children's peer relationships. This is quite a curious outlook: *The peer group needs greater freedom to work out its problems than most other parents are willing to delegate to it.* It suggests that our informants put an early and strong emphasis on independence training with respect to handling peer relationships and might be relatively indifferent to aggressive skirmishes within the peer group itself. Unfortunately, there are not enough items in the battery of situations to document fully the parental philosophy toward their children's peer groups, but the general findings deserve careful consideration in further studies of this sort.

The last sub-group consists of behaviors relating to *oral and dependency conflicts* of children—i.e., not wanting mother to leave, sucks thumb, bites nails, cries easily, etc. These mothers try to treat such oral dependency problems and general manifestations of “nervousness” with more permissiveness than most neighbors are judged to display. It is as if our informants believed that most parents fail to recognize that nothing is to be gained by punitiveness in coping with oral dependency problems; only the nurturance and comforting attention involved in the use of “sweet reason,” or the security and freedom involved in allowing the child to master these problems without signs of parental displeasure or anxiety, are really felt to work pragmatically.

It is curious that enuresis is handled so entirely differently from daytime loss of bladder-control. The latter was treated very much like disobedience and aggression, whereas the former is treated more like a dependency problem. The disparity might merely be a function of the degree of public

shame involved in daytime loss of bladder control, which reflects on the status of the parents. However, it might also represent a tendency for norms governing toilet-training to vary more from one toilet-training situation and problem to another. To put it differently, the norms regarding toilet-training may be less general than those governing other areas of socialization, especially the secondary and culturally acquired drives of aggression and anxiety. The lack of generality of norms for toilet-training, and perhaps for most of the biological drives, might be taken to indicate either a greater inconsistency in socialization or, instead, a willingness to pay more attention to the special circumstances under which each specific kind of toilet-training arises rather than to adopt a blanket, wholesale discipline policy regarding toilet matters. Our data provide no way to decide among the alternatives. However, there is generally thought to be more intense anal interests among individuals in working-class populations and, according to Whiting's theory of acquired drive as reported in Sears (7), one would anticipate that inconsistency of training would accentuate strong motivational interests in toilet matters.

Summarizing these results it appears that external and internal norms converge with respect to sanctions and taboos regarding two specific areas of child training. Fears and anxieties of children are to be handled non-punitively, but the defiance of the parents' authority by the child gives the parent a right to resort to punitive measures. The disagreements between external and internal norms, however, cover a greater variety of situations and problems. Oral dependency problems, the child's relationship to peers, and the reluctance of the child to do what is "naturally" good for him are all grounds for punitive measures according to external norms. However, the mothers report themselves as having been more influenced in their actual treatment of such problems by the internal norms which taboos punishment. The fact that the mothers depart from the external norm in a greater variety of situations than they conform to it may be taken as additional evidence against the other-directedness of our subjects, and such an interpretation is consistent with the rest of the findings.

D. DISCUSSION

Using parental norms of discipline, the attempt to test the hypothesis that there is a general social tendency toward other-directedness has provided some strong evidence against it. The ranking test of discipline methods demonstrated that, when internal and external ranking-norms diverge, the mothers were strongly inner-directed. Furthermore, the external norm

did not possess as clear-cut or compelling discipline sanctions as the internal norm, nor did it reflect as organized and general a philosophy of child care. When a subject-by-subject categorization was made into two normtypes, only one mother out of the 30 could be categorized as other-directed. The mothers, furthermore, deviated from the external norm in a more uniform and consistent direction than they did from the internal norm indicating a common, shared, and critical attitude toward neighbors, incompatible with an other-directed approach. And, lastly, the situations in which the mothers perceived themselves as acting in accordance with their neighbors were of a quite specific nature, whereas they perceived themselves acting differently from their neighbors, and in terms of their internal standards, in a much wider variety of situations. The evidence dovetails, therefore, into a highly self-consistent picture of the inner-directedness of these mothers.

The Riesman theory, however, is meant to apply to many other social rôles than those of parents, to many other sources of external subtle kinds of group influence than we have set out to measure. The need to introduce further complexities in future research of this kind do not diminish the value of our data. Rather, the data allow us to raise some interesting questions as to which complications are to be introduced (and why) at the simplest level, which might have been missed by more complex and complicated initial research undertakings. These questions involve seven ambiguities in Riesman's theory.

It is not impossible that a good deal of what Riesman has included under the rubric of "other-direction" is the capacity to empathize, which allows one to continue to learn from others, and need not represent a total collapse of the guiding function of internal standards and independent thought. Upper-middle class suburbanites, unlike our housing-project mothers, might be found to show a preponderant other-direction had we studied them. But a question would still remain whether, with greater economic security, and better educational opportunities, people would be more willing to have their favorite views and principles modified, rather than continue to conform rigidly to standards whose rightness is never questioned, and which perhaps persist as exact replicas of the early models from which they were first introjected. Furthermore, the modification of their ideals may be in terms of rational, independent deliberation of the advantages offered by the various models around them, rather than blindly and slavishly. The mothers in our study appear to have little of this kind of empathy, but instead their "inner-direction" veers toward a belligerent hostility, perhaps even a kind of guilt and projection. Indeed, as one reads Riesman, both inner-

direction and other-direction seem equally unpleasant phenomena. The type of personality that is able to integrate, by the effective and sometimes painful use of intelligence, wisdom introjected from the past with wisdom occurring in the contemporary scene, could easily be confused with "other-direction." It would be a sad thing if our conceptual tools did not allow us to differentiate a thoughtful and warm receptiveness toward others from a wholesale, indiscriminating fashionableness, and a non-messianic, informed, independent kind of thinking from a know-it-all, aggressive self-righteousness. Yet, unnecessary ambiguities are bound to arise, and with them possible excesses of statement and interpretation, which could be lessened by making these differences in receptiveness to others more explicit than they have been made.

There are still further sources of ambiguity, which more specifically impede research study. Each mother in our study may see reflected around her in her neighbor's behavior three or four different standards of parental discipline. Her internal norm may have a quite definite reference point, whereas the external norm is somewhat heterogeneous. The very heterogeneity of the external norm, as seen by our informants, may therefore make her adhere more to her internal standards in verbal reports. At the same time, each mother may be heavily, but selectively, influenced by some two or three mothers in her immediate environment, who share homogeneous norms, and *even more so than by her own norms*, whereas she remains "inner-directed" with respect to *most* of her neighbors. How can one specify whether such a mother is "inner-directed" or "other-directed" without specifying toward whom the term refers? It may be that, had we asked our mothers to report what practices mothers whom they *admired most* employed in disciplining their children, we would have found marked other-directed tendencies. But, in doing so, we would be introducing a dimension into the concept of "other-direction" which is not altogether consistent with some of its meaning. There is a need, in other words, to remove the ambiguity which allows us to place both a *discriminating* other-direction, based on whatever criteria, and the tendency *to be influenced by everyone* in the same category. As the Riesman hypothesis stands, it cannot help but prove true, although not very helpful, if we but continue to search for the right kind of "other" which fulfills the idea. In this way the meaning of "other-direction" could change from class to class, from rôle to rôle, and from individual to individual without being noticed. There could be no refutation of a theory, used this way, by any amount of empirical research.

It is conceivable that as problems of this sort are clarified we will begin

to find that individuals and societies cannot be categorically classified as either inner-directed or other-directed. Instead, we may begin to discover that all individuals and societies are in some sense both, depending on the area of personality under consideration and the type of "other" being considered. Periods of high insecurity in the life cycle may give rise to now one, now the other, kind of extreme tendency, depending on how growth can best take place, which are preceded by and pass into high levels of integration of internal and external norms. In other words the tendencies toward either kind of norm conformity may be a normal, non-pathological *stage* in a growth process, which *recur* in either direction throughout the life cycle around special life problems and special relationships which have still to be resolved. As norm integration in one area of life and in one type of social relationship is resolved, new norm conflicts may develop on the frontiers of still further social learning and experience. Such a scheme is as easily applicable to groups as it is to individuals. Only an individual or society in which no long-term social growth or learning took place could be classified as categorically inner-directed or other-directed. It is hard to tell whether this kind of image of alternating and partial norm integration and conflict is consistent with Riesman's views about conformity, or whether it diverges markedly from them, since so much of his thinking depends upon the use of an "ideal type" method of analysis. It appears to us to do equal justice to the kind of factual materials he has brought together in his writings and throws a totally different light on the data within this study.

It may be argued that had we measured the *real*, rather than the *reported*, behavior of each mother, other-directed tendencies would have become more apparent among these mothers. In such an event, the hypothetical tendency toward other-direction would have to be looked upon as an *unconscious tendency* and measured by some other means than those depending upon conscious reporting. It does not seem to us that Riesman has particularly emphasized the unconscious nature of other-directedness in his writings, and there are many passages which seem clearly to imply a self-conscious kind of reinforcement process, of at least marginal awareness, for behaving in some ways consistent with "the crowd." It seems to us that *only if* "other-directed" tendencies are to be considered unconscious, suppressed or repressed because of shame or guilt about them, does any serious matter arise regarding the validation of the mothers' reports arise. Otherwise, there would be no motivation which would require *systematic* distortion and misperception of tendencies toward other-direction. However, apart from the bearing of this issue on the data of our study, it is worth considering in its

own right. Unconscious tendencies are acquired, function, and become modified in different ways than conscious tendencies. Furthermore, "insight" often develops regarding various unconscious tendencies during the course of the life cycle. If, therefore, the concept of other-direction is to be understood as being capable of an unconscious presence, it is very important to distinguish the operation of this tendency, and its opposite, under the two conditions of consciousness. It is especially important to distinguish the concept of "unconscious other-direction" from "suggestibility," since it seems relatively clear from a variety of studies that little evidence can be found for any tendency toward *general* suggestibility. And, lastly, there is the problem of understanding why other-directed tendencies should be repressed or suppressed in an other-directed society which presumably encourages them.

Nevertheless, our informants do appear to be poor reporters, both regarding their own and their neighbors' behavior, for their pooled reports hardly stack up with great internal consistency against one another. This might be the result of a hostile projection, whereby mothers report themselves as following their own standards, which in reality they may not do, and project their felt inadequacies onto other mothers in the environment. It might also be the result of each mother only being able to observe her neighbors disciplining their children in public situations, and then, extrapolating from these incorrectly to discipline situations which are confined to intimate family interactions. Some of the discrepancy could be explained this way, but not all, for it holds up especially well for matters pertaining to the peer group, whose supervision is likely to be most public of all. It is possible that the *actual* behavior of these mothers does follow the norms they attribute to their neighbors, although this norm is the result of a projection, made possible by the fact that it is established around situations which to some unmeasured degree are removed from prolonged first-hand observation. In such a case, a very marked change of meaning would accompany the categorization of such mothers as "other-directed," for the tendency is now the result of a projection which has itself been instigated by the guilt arising from the infraction of a superego norm. To put the matter concisely, to what degree does the external norm have to be "realistic," based on accurate and possible observations of the behavior of others, before an individual who follows it can be characterized as other-directed?

In this regard, it is relatively easy to see how an other-directed individual is socially influenced by an inner-directed person, since the latter presents salient social cues. But it is much more difficult to understand how two

other-directed persons influence one another at all, since both tend to look for social cues from one another. Where an entire social group consists of such individuals, the process of social influence is harrowing to contemplate, for it is hard to discover how any important action gets initiated or completed without overwhelming anxiety on everyone's part. Without having a sketch of how standards arise in such groups, and what enables some types of other-directed personalities to conquer their anxieties, so as to provide leadership, the concept has little usefulness. The process of "marginal differentiation" is applicable only *after* a standard has been initiated and agreed upon. The problem is how other-directed groups ever come to agree upon standards at all.

The term "other-directed" connotes a process of social influence but it is also many times used to refer to special kinds of beliefs and behaviors, quite apart from the social source of their derivation. Thus, our mothers, because they appear to believe the peer-group should be allowed to work things out for themselves and because they believe in discussion as an ideal discipline method, might be characterized as other-directed, even though they do not perceive themselves getting such ideas from one another. In Riesman's writing, for example, it appears that parents who believe that there are no problem children, only problem parents, should be characterized as other-directed without having to take into account how they have come to hold such views. They may have interiorized such a belief from their parents and primary socializers, or they may have picked up a fashionable idea from their club or neighborhood interactions, or as still another possibility, they may have arrived at the belief by highly painful and creative thinking and introspection. It is not at all clear, since so many avenues may lead to the same beliefs, why the belief should be characterized as intrinsically "other-directed" without taking into account the conformity process. When we consider the immense differences there would be among such mothers in attempting to modify their views, the issue becomes even more salient. It would be difficult to reason with the first mother because of unconscious identification, whereas the third mother might be fairly soon reasoned out of, as she was into, her belief by carefully examining her arguments and insights and correcting them. The second mother, through a change of neighborhood or of her mass-media diet, might quickly be brought around to give up her belief. It is only this second mother which seems other-directed to us. She would presumably be still as other-directed when the fashion suddenly changed, once we know how they do change in an other-directed society, and she begins to believe that there are no problem parents, only problem

children or some other variation on the theme. The example serves, in any event, to make clear one difficulty in trying to extend the term other-direction to include certain beliefs and behaviors apart from the way in which they were socially acquired or can be socially modified. When this kind of extension of meanings is made, no amount of empirical research can ever refute an other-directed hypothesis, for what does not seem to be influenced by others may still be viewed as intrinsically part of the syndrome. The prophecy is, by the very nature of the logic, bound to be self-fulfilling.

There is a way out of this kind of self-containment, however. It may be that certain types of beliefs and practices are *usually* or *most frequently* the result of a process of conforming with external norms. But this, then, becomes an empirical matter, not something to be decided in terms of the intrinsic nature of the belief or practice itself. Each such indicator, regarding special beliefs and behaviors, has to be correlated with the social process by which they were acquired.

Otherwise, the concept of "other-directed" has so many meanings that it loses any precise meaning at all. The correlation of such indicators with one another, in the absence of some sort of independent test regarding their social derivation, merely allows us to say that certain beliefs go together as we had predicted. They could go together, however, for all sorts of reasons, only one of which is that they are all the result of "other-direction." For instance, they might intercorrelate because the individual has thought through, by himself and in a consistent manner, his social philosophy, or on the other hand and much worse, because he has not thought about them at all and has had no occasion to do so and blindly agrees out of thoughtlessness, ignorance, or indirect prestige-suggestibility. In a private communication, Herbert Kelman has reported that he finds inner-directed and other-directed subjects, defined in terms of agreement with such appropriate statements of belief indicated by Riesman, *both* tend to score high on tests of authoritarianism. The finding serves to accentuate the dangers of *assuming* a set of beliefs are acquired in a certain kind of way merely because they intercorrelate.

A group of problems involving seven ambiguities have thus come out of our attempt to relate our findings to the writings of David Riesman. Briefly summarized for the convenience of the reader, these are:

1. To what degree, how and under what conditions may internal and external norms be integrated rather than given preferential status over one another?
2. How indiscriminating does susceptibility to influence by others have

to be before it can be characterized as "other-directed" and what criteria, if any, of discrimination are used by the other-directed?

3. To what degree may individuals or societies be characterized as *partially* inner-directed or other-directed by viewing them as recurrently within a learning situation involving the partial progressive integration of external and internal norms?

4. To what degree and under what conditions may other-direction or inner-direction be considered an unconscious process, why does this happen, and does the unconscious process differ from its conscious correlate in modification and functioning?

5. To what degree does the external norm have to be "realistic" before influence by it is characterized as other-direction?

6. Does the process of other-direction differ in a group composed entirely of other-directed people from one which is mixed, and if so, how does the former kind of group come to establish norms?

7. Does the concept of other-direction apply to sets of beliefs and practices, independently of how they are personally acquired, and if so, does the concept not then begin to shift its meaning radically?

While these seven ambiguities arose out of our experience in conducting this particular study and relating it to the theory, they will undoubtedly arise again, in one form or another, in any further empirical work with the hypothesis of other-direction and on that account alone deserve further analysis. Taken together with their unambiguous counterparts in the writings of Riesman, they serve to define a useful area of research which if followed through, could extend our knowledge of conformity behavior even beyond the advanced level to which Riesman has already brought it.

The techniques used in this study are capable of modifications in ever more subtle ways but, even as they stand, they allow us to get some initial picture of areas in which general norms tend to break down and shift in character, as well as providing measures of self-consistency within any one area of interpersonal relations. There are some special problems arising from the data which deserve special mention and consideration aside from those already mentioned—i.e., socio-economic differences in child training norms, and differences in the consistency with which primary and secondary motivational systems are handled by the parent.

None of the mothers in our studies used religious appeals in socializing children. Yet religious affiliations have become stronger, not weaker, since the last war. It hardly seems reasonable, then, to assume that religious appeals are not used by parents to socialize their children. Rather, it seems

likely that the special religious appeal we selected is not that most frequently preferred or used by religious parents, and additionally, the situations which we chose were not those in which religious appeals would be resorted to by parents. Further research has been planned to get answers to these questions and to test some hypotheses about how and under what circumstances religious ideology enters into child training.

Secondly, the area of peer relationships seems, among these mothers, to be one in which they feel themselves markedly at odds with other parents. According to Parsons (4), peer relationships should take on high importance as an avenue of linking the family with the general community, since the American family is comparatively more insular than it is in other societies. Yet, it seems to be an area in which these parents feel themselves most in disagreement and conflict with each other concerning norms of socialization. Such a perception cannot help but further reinforce the insularity of the parent which Parsons depicts and, simultaneously, the insularity of the peer-group from the families of its members. To what degree such an emphasis on independence training regarding peer-group relationships is a form of rejection of the child, creating an atmosphere in which juvenile delinquency might take hold, is still an open question. Certainly, one would not expect the same results would be obtained in the new suburbias, where it appears from sociological studies that peer groups function instrumentally to bind families together. To the extent that insularity is similar to inner-direction, our study tends to support the picture of the American family that Parsons has outlined.

E. SUMMARY

An attempt was made to discover the relative strength of "inner-directed" and "other-directed" tendencies in a group of mothers relatively low in the social strata. Four kinds of test provided a self-consistent picture that the informants were more inner-directed than other-directed in their discipline procedures used in socializing children. It was also found that discussion was the general norm of discipline among the mothers, except when the child aggressed against his own family, or extensions of it, quite openly and directly. In the latter event the discipline norms of the parents shift toward a more punitive frame of reference. Seven ambiguities regarding the meaning of Riesman's terms were discussed as they related to the data and further uses of the techniques employed in the study were suggested.

REFERENCES

1. BAUER, R. A. The psychology of the Soviet middle elite: Two case histories. In Kluckhohn, C., Murray, H., & Schneider, D. M. (Eds.), *Personality in Nature, Society, and Culture*. New York: Knopf, 1954.
2. BRODBECK, A. J. A conceptual analysis of the Riesman typology. (In press.)
3. LEWIS, O. *Life in a Mexican Village: Tepoztlan Restudied*. Urbana, Ill.: Univ. Illinois Press, 1951.
4. PARSONS, T. *Essays in Sociological Theory*. Glencoe, Ill.: Free Press, 1954.
5. REDFIELD, R. *A Village That Chose Progress: Chan Kom Revisited*. Chicago: Univ. Chicago Press, 1950.
6. RIESMAN, D. *The Lonely Crowd*. New Haven: Yale Univ. Press, 1950.
7. SEARS, R. R. Ordinal position in the family as a psychological variable. *Amer. Sociol. Rev.*, 1950, **15**, 397-401.
8. STENDLER, C. B. The impact of beginning first grade upon socialization as reported by mothers. *Child Devel.*, 1950, **21**, No. 4.

*Center for Advanced Study in the
Behavioral Sciences
202 Junipero Serra Boulevard
Stanford, California*



THE RELATION BETWEEN LEARNING ABILITY SCORES OBTAINED IN DIFFERENT SITUATIONS*¹

*Institute of Child Welfare, University of Minnesota, and USAF School
of Aviation Medicine*

MERRILL ROFF AND R. B. PAYNE

The problem of the generality or specificity of learning ability, when different learning tasks are employed, is of critical importance in determining whether a learning ability score obtained from any one task can be considered in any way a general characteristic of the individuals involved. Learning is normally defined as a change in performance due to training, experience, etc.; this is also the definition implied by all treatments of learning in terms of learning curves, which present pictorially a change from initial to subsequent status. Correspondingly, a learning ability score is the difference between initial and final performance on some learning task.

The main problem considered in the present paper is that of the relation between learning or gain or improvement scores on two or more different tasks. Following what seems to have been the initial study in this area by Wimms (15), who determined the correlation between "improvability" on two different computational tasks with a small sample of teen-age boys, the first integrated attack on this problem was a series of studies by Thorndike (13) and his associates, Chapman (3) and Race (9). A later group of interrelated studies have been described by Woodrow (16, 17). Explicit data on this problem have also been presented by Gates (7), Brooks (1), Hall (8), and Roff, Payne, and Moore (10).

A special case of this problem is that in which a single measure (intelligence quotient, aptitude test, etc.) is related to change of performance on a second variable. Some of the studies cited above also present data on this point; additional reports of such data have been given by Ewert (6), Drake (4), Dysinger and Gregory (5), Carlson, Fisher, and Young (2), Simrall (11), and Tilton (14).

All these studies vary widely in samples studied, in learning tasks employed, in amounts of practice given, and in other situational factors; they

*Received in the Editorial Office on September 1, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹This study was supported (in part) by funds provided under Contract AF 18 (600)-319 with the USAF School of Aviation Medicine, Randolph Field, Texas.

also vary widely in quantitative results obtained. One of the purposes of the present paper is to indicate that there is no sound basis for expecting that results obtained on problems such as this in any one learning situation will correspond at all closely with results to be expected in other, and different, situations.

The derivation of learning ability scores as the difference between initial and final performance on some task permits the development of a statistical procedure whereby an exact determination can be made of the correlation between the gain or improvement in any two different learning situations, as a function of parameters of the initial and final performances on the tasks employed. This is advantageous, since much more is known about the correlations among performance scores on a wide variety of tasks than is known about the correlations among their learning scores. This approach gives more nearly general answers to problems in this area than can be obtained by the analysis of any single set of experimental data, or by the detailed collation of previous research results, by indicating the relations to be expected between gain scores generally, when information is available about initial and final performance measures in different learning situations.

The formula for the correlation between gain or improvement in one task, and gain or improvement in a second task, can be written explicitly in terms of the correlations between initial and final trials in the two situations, and the standard deviations of the two initial trials and the two final trials, as follows:

$$r_{g_1 g_2} = \frac{r_{f_1 f_2} \sigma_{f_1} \sigma_{f_2} + r_{i_1 i_2} \sigma_{i_1} \sigma_{i_2} - r_{f_1 i_2} \sigma_{f_1} \sigma_{i_2} - r_{i_1 f_2} \sigma_{i_1} \sigma_{f_2}}{\sqrt{(\sigma_{f_1}^2 + \sigma_{i_1}^2 - 2 r_{f_1 i_1} \sigma_{f_1} \sigma_{i_1})(\sigma_{f_2}^2 + \sigma_{i_2}^2 - 2 r_{f_2 i_2} \sigma_{f_2} \sigma_{i_2})}} \quad (1)$$

Where g_1 = gain on Task 1
 g_2 = gain on Task 2
 f_1 = final score on Variable 1
 f_2 = final score on Variable 2
 i_1 = initial score on Variable 1
 i_2 = initial score on Variable 2

Consideration of the numerator of this expression by itself immediately indicates one definite conclusion:

If two tasks administered to some sample show no correlation among initial and final trials, the correlation between their learning scores will also be zero. With such tasks there is no possibility of deriving a learning score that can be considered a general characteristic of the individuals involved.

Less rigorously, the conclusion is also indicated that if the correlations among initial and final trials are near zero, correlations among learning scores will also be near zero, so that there is no possibility of obtaining a learning score from any one task which will have enough generality for the individuals concerned to be of any practical significance. Several of the studies mentioned above include at least one pair of tasks to which these conditions apply.

Before discussing other results indicated by this formula, it is desirable to consider the relative size of the initial and final standard deviations for a single learning task, since these values contribute to the determination of the values of the dependent variable. This problem may also be stated as follows: does practice on a task make a group more or less variable than they were at the beginning?

This problem is inseparable from a second one: what is the correlation between initial performance on a task and the improvement which occurs on that task as a result of practice?

The kind of answer to the first of these questions which will be obtained in any empirical study is given directly by the formula for the variance of a sum of two scores, since the final score for each person is by definition the sum of his initial score and his learning or gain score,

$$\sigma_t^2 = \sigma_i^2 + \sigma_g^2 + 2r_{ig} \sigma_i \sigma_g \quad (2)$$

If the standard deviation of initial scores is set equal to 1.00, as can always be done, this expression becomes

$$\sigma_t^2 = 1 + \sigma_g^2 + 2r_{ig} \sigma_g \quad (3)$$

From this it is easily possible to construct a table showing the values of the standard deviations of final scores which will result from various combinations of standard deviations of gain scores and correlations of initial and gain scores (Table 1).

It is immediately apparent from Table 1 that there is no single answer to the question of whether variability of performance increases or decreases with practice. For some tasks administered under some practice conditions to some samples of persons, learning may result in an increase in variability of performance, while for other tasks in relation to the same or other samples of persons, variability may decrease, or remain about the same. It may be noted that a decrease of variability occurs only in relation to negative correlations between initial and gain scores.

The correlation between initial performance and gain on a task can be

indicated by solving Equation 2 explicitly for the correlational term. It then appears as a function of the standard deviations of the initial, gain, and final scores. Since these have different relative values for various tasks and samples of persons, there is no single true value for the correlation between ini-

TABLE 1
STANDARD DEVIATIONS OF FINAL SCORES AS A FUNCTION OF THE STANDARD DEVIATIONS OF, AND CORRELATION BETWEEN, INITIAL AND GAIN SCORES, WHEN THE INITIAL STANDARD DEVIATION IS 1.00

		STANDARD DEVIATION IS 1.00							
		Standard deviation of gain scores							
		.10	.30	.50	.70	.90	1.10	1.30	1.50
Correlation between initial and gain scores	— .80	.92	.78	.67	.61	.61	.67	.78	.92
	— .60	.94	.85	.81	.81	.85	.94	1.06	1.20
	— .40	.96	.92	.92	.96	1.04	1.15	1.28	1.43
	— .20	.98	.98	1.02	1.10	1.20	1.33	1.47	1.63
	.00	1.00	1.04	1.12	1.22	1.35	1.49	1.64	1.80
	.20	1.02	1.10	1.20	1.33	1.47	1.63	1.79	1.96
	.40	1.04	1.15	1.28	1.43	1.59	1.76	1.93	2.11
	.60	1.06	1.20	1.36	1.53	1.70	1.88	2.06	2.25
	.80	1.08	1.25	1.43	1.62	1.80	1.99	2.18	2.38

tial performance on tasks and the amount of improvement which occurs on these tasks as a result of practice. In some situations this correlation may be positive, and in others markedly negative; values in the literature cover a wide range, depending on the tasks used and the samples studied. The effect of error of measurement on these correlations (12) is minor in comparison with the effects of all the other situational factors involved.

One particular combination of circumstances has occurred or has been approximated in a sufficient number of studies to make it worth discussing in some detail. This is the situation in which the final task following learning does not have enough performance ceiling to permit much variability of final performance. Let us assume, as a limiting case, a task which could be perfectly performed by all members of a group if enough practice or training were given. In such a case, variability of final performance would be zero, and the correlation between initial and gain scores would be -1.00 , since those initially lowest would eventually make the largest gains. One consequence of such a situation is that any aptitude test which correlated positively with level of initial performance would correlate negatively with the amount of gain or improvement with practice or training. We thus have the superficially paradoxical result that aptitude or intelligence is negatively

correlated with "ability to learn." A different result would of course appear if the experimental situation were changed to permit an increase of variability in final scores.

Returning now to the original problem of the correlation between improvement or gains on two different tasks, comparable tabulations can be made for Formula 1 above. However, since this expression contains more variables than Formula 2, which relates to one task only, such a tabulation can be made only by assuming certain values as constant and letting other values vary, to find what the effect of such variation will be on the correlation between the two improvement scores. Thus, one set of constants, which approximates somewhat some available research results, has been assumed for the computation of Table 2. The standard deviations of initial scores on

TABLE 2
CORRELATIONS BETWEEN GAINS ON TWO TASKS, ASSUMING:

$$\sigma_1 = \sigma_1 = 1.00,$$

1 2

$$\sigma_g = \sigma_g = 1.00,$$

1 2

$$r_{1g} = r_{1g} = .00,$$

1 1 2 2

$$r_{1f} = r_{f1}, \text{ and}$$

1 2 1 2

$$r_{1f} = r_{f1}.$$

1 2 2 1

		Correlation: initial-initial and final-final									
		.00	.10	.20	.30	.40	.50	.60	.70	.80	.90
Correlation: initial, each test with final of other	.00	.00	.30	.60	.90						
	.10	— .28	.02	.32	.62	.92					
	.20	— .57	— .27	.03	.33	.63	.93				
	.30	— .85	— .55	— .25	.05	.35	.65	.95			
	.40		— .83	— .53	— .23	.07	.37	.67	.97		
	.50			— .81	— .51	— .21	.09	.39	.69	.99	
	.60				— .80	— .50	— .20	.10	.40	.70	1.00
	.70					— .78	— .48	— .18	.12	.42	.72
	.80						— .76	— .46	— .16	.14	.44
	.90							— .75	— .45	— .15	.15

the two tasks are taken as 1.00, which can always be done. It has been assumed that the standard deviations of the two final scores have a value of 1.41; this is the value which would result if the standard deviation of gain scores in each task is also 1.00, and the gain scores are uncorrelated with their own initial scores in each case. This is a situation in which the variability of both tasks shows some increase with practice. If we further assume

that the correlation between the two tasks remains constant with practice, as it not infrequently does, so that the correlation between initial performance on the two tasks equals the correlation between the two final performances, and finally that the correlations $r_{1\frac{f}{2}}$ and $r_{2\frac{f}{1}}$ are equal, the correlation be-

tween gain scores on the two tasks can be read directly from the table for various levels of correlation between the two tasks. Table 2 shows the various correlations between gain scores which will result with these assumed values. It should be realized that the multiplicity of assumptions made here is in no way essential to the application of Formula 1; their use is simply an expedient to permit a tabulation of this type. Additional tabulations based on different assumptions have been presented elsewhere (10).

Inspection of Table 2 indicates that with these particular constant values, if the four correlations involved are approximately equal, the correlations of gains or improvement on the two tasks will tend to be near zero. If the correlations between the two initial and the two final performances exceed the two correlations between initial scores on one task and final scores on the other, the correlation of gain scores will be positive, and sometimes substantially so. If this latter situation is reversed, the correlation between gain scores will be negative. Obviously, no single cell in the table can be taken as representing the correlation between improvement scores for learning in general. It is not possible to make an accurate inference even as to the sign of the correlation between improvement scores from one pair of tasks to another quite different pair of tasks; the results are specific to the experimental situations and samples studied. On the other hand, sets of markedly different tasks may happen to yield similar results.

Whether or not it would be possible to find or arrange situations which would yield the results indicated in each cell of this table is not known. It would probably be necessary to have errorless measures to obtain some of the more extreme positive or negative values. In two previous studies (3, 16), the highest positive correlations between gains were found for pairs of tasks so similar that transfer from one to the other may have occurred. Whether this is a necessary condition for high positive correlation between gains is not known.

Taken all together, the considerations of the present paper raise rather sharply the question as to what rationale there is for expecting that specific quantitative results obtained with any one task on any one specific sample of organisms can be extended with confidence to completely different tasks administered to other samples of organisms, which differ markedly in meth-

ods of selection from the original sample. The standard procedures of statistical inference are, of course, irrelevant in this connection, since they are based on a continuing identity of both experimental conditions and of method of selection of individuals, from one sample to another. This problem would seem to be particularly acute when, as is sometimes done, information is sought concerning learning in humans by studying other animal species, which may be regarded as "simpler," more amenable to experimental manipulation, etc. In such a situation there would seem to be no rationale for expecting that quantitative results of learning experiments would hold in any significant detail for different tasks, employed with any specific sample of humans (at any age level, etc.).

If practical applications of the results of research on problems of learning are intended, the implication seems clear that the samples of persons studied and the conditions of the learning situations employed should be as similar as possible to the practical situation in which the information is to be applied. Lack of similarity of subjects or of situation between an experimental study and the point of desired application would seem likely to lessen markedly the likelihood of correspondence between the results of experimental studies and outcome in practical situations.

REFERENCES

1. BROOKS, F. D. Learning in the case of three dissimilar mental functions. *J. Exp. Psychol.*, 1924, 7, 462-468.
2. CARLSON, H. B., FISHER, R. P., & YOUNG, P. T. Improvement in elementary psychology as related to intelligence. *Psychol. Bull.*, 1945, 42, 27-36.
3. CHAPMAN, J. C. Individual differences in ability and improvement and their correlations. *Teach. Coll. Contr. Educ.*, 1914, No. 63.
4. DRAKE, C. A. The Iota function. *J. Educ. Res.*, 1940, 34, 190-198.
5. DYSINGER, D. W., & GREGORY, W. S. A preliminary study of some factors related to student achievement and grades in a beginning course in psychology. *J. Gen. Psychol.*, 1941, 24, 195-209.
6. EWERT, H. The effect of practice on individual differences when studied with measurements weighted for difficulty. *J. Gen. Psychol.*, 1934, 10, 249-285.
7. GATES, G. S. Individual differences as affected by practice. *Arch. of Psychol.*, 1922, No. 58.
8. HALL, C. S. Intercorrelations of measures of human learning. *Psychol. Rev.*, 1936, 43, 179-195.
9. RACE, H. V. Improvability, its intercorrelations and its relations to initial ability. *Teach. Coll. Contr. Educ.*, 1922, No. 124.
10. ROFF, M., PAYNE, R. B., & MOORE, E. W. A statistical analysis of the parameters of motor learning. USAF School of Aviation Medicine, Project No. 21-0202-0001, Report No. 1, Feb., 1954.
11. SIMRALL, D. The effects of practice on the factorial equations for visual, spatial, and perceptual tests. Ph.D. Dissertation, Univ. Illinois, 1946.

12. THOMSON, G. H. A formula to correct for the effect of errors of measurement on the correlation of initial values with gains. *J. Exp. Psychol.*, 1924, 7, 321-324.
13. THORNDIKE, E. L. Notes on practice, improvability, and the curve of work. *Amer. J. Psychol.*, 1916, 27, 550-565.
14. TILTON, J. W. Intelligence test scores as indicative of ability to learn. *Educ. Psychol. Measmt.*, 1949, 9, 291-296.
15. WIMMS, J. H. The relative effects of fatigue and practice produced by different kinds of mental work. *Brit. J. Psychol.*, 1906, 2, 153-195.
16. WOODROW, H. Interrelations of measures of learning. *J. of Psychol.*, 1940, 10, 49-73.
17. ———. The ability to learn. *Psychol. Rev.*, 1946, 53, 147-158.

Institute of Child Welfare
University of Minnesota
Minneapolis 14, Minnesota

Department of Psychology
USAF School of Aviation Medicine
Randolph Air Force Base, Texas

A NEW ILLUSION OF APPARENT MOVEMENT AND THE CONCEPT OF RETINAL LOCAL SIGNS^{*1}

Department of Psychology, Western Reserve University, and the Highland View Hospital, Cleveland

JAN H. BRUELL AND GEORGE W. ALBEE

A. PROBLEM

We would like to present a brief report on a compelling new illusion of apparent movement, together with a description of the considerations which led to its discovery. In the discussion, we will mention some implications of the phenomenon for a theory of retinal local signs.

If, in a dark room, we place a luminous vertical line in the objective median plane of an observer, he will see it as straight ahead. If, however, the same line forms the vertical edge of a luminous area, it will not be seen as straight ahead. Roelofs (3), for example, found that if a vertical line forms the right edge of a rectangle, an observer fixating that edge will see the rectangle at a distance X to his right in spite of the fact that the fixated edge is placed objectively straight ahead of him. (Although the entire figure is displaced, for convenience of exposition we will refer henceforth only to the displacement of the fixated edge of the stimulus.)

In an elaboration of Roelofs' experiment we have found (2) that the magnitude of this perceptual displacement of the fixated edge of a luminous stimulus depends on the horizontal extent of that stimulus. In the example above, if the rectangle extends a distance A to the left of the fixated edge, that edge will be displaced perceptually a distance X to the right. If the rectangle extends a distance $2A$ to the left of fixation, the fixated right edge will be displaced a distance $2X$ to the right. These perceptual displacements are immediate. As soon as the stimulus is presented, its right edge is seen a distance X or $2X$, etc., to the right.

These data suggested the following experiment. If an observer fixated the right edge of a very narrow upright rectangle placed in his objective median plane, he would see this edge as straight ahead. Now let us expand the

^{*}Received in the Editorial Office on September 2, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹From the Department of Psychology, Western Reserve University, and the Department of Physical Medicine and Rehabilitation, Highland View Hospital.

The experimental work reported was supported by a grant from the National Science Foundation.

narrow rectangle gradually into a rectangle with a horizontal extent A , $2A$, etc., by adding luminous area to the edge opposite the fixated edge. The fixated right edge of such an expanding area should appear to move a distance X , $2X$, to the right. Conversely, an observer fixating the left edge of a luminous rectangle expanding slowly to the right should experience apparent movement of the fixated left edge toward the left. In other words, it was hypothesized that while in Roelofs' experiment the fixated edge of a rectangle appeared displaced on presentation, in such a situation as we have described the displacement of the fixated edge from a medial to a lateral position should become visible *in statu nascendi*.

B. METHOD

The experiment was performed in a dark room. Twenty-three Ss, undergraduates at W.R.U., were used. All Ss were tested individually. They were seated in a chair provided with a headrest which kept the S's head in position. The test figure, as S first saw it, was a luminous vertical bar 20 cms. high and 2 cms. wide. The S was asked to fixate one edge of the bar and to continue fixating this designated edge throughout the trial. He was asked also to report anything he saw during the trial. Care was taken in the instructions not to suggest the possibility of movement of the fixated edge.

When S understood what he was to do and indicated that he was fixating the designated edge of the bar, E turned on a motor which slowly pulled a masking panel in a horizontal direction exposing a luminous area which gradually increased in horizontal extent (see Figure 1). The speed of exposure was 2 cms. per second until a rectangle 40 cms. wide was exposed. Each S was tested under two experimental conditions. One of the conditions required S to fixate the left edge of the bar while the masking panel was pulled to the right, thus slowly forming a rectangular luminous area extending to the right of fixation. With the second condition, the right edge of the bar was fixed and a rectangular luminous area to the left of fixation was uncovered slowly. Figure 1 illustrates the latter condition. A verbatim record was made of S 's report during testing.

C. RESULTS

All but one S reported movement of the fixated edge in the expected direction under at least one of the two experimental conditions. When the right edge of the luminous bar was fixated and the area to the left of fixation was uncovered, 18 of the 23 Ss spontaneously reported movement of the

fixated edge to the right, that is, in the predicted direction. Five *Ss* did not report movement. No *S* reported movement to the left.

Under the opposite condition, 20 of the 23 *Ss* spontaneously reported movement of the fixated edge to the left. Three *Ss* did not report movement. No *S* reported movement to the right.

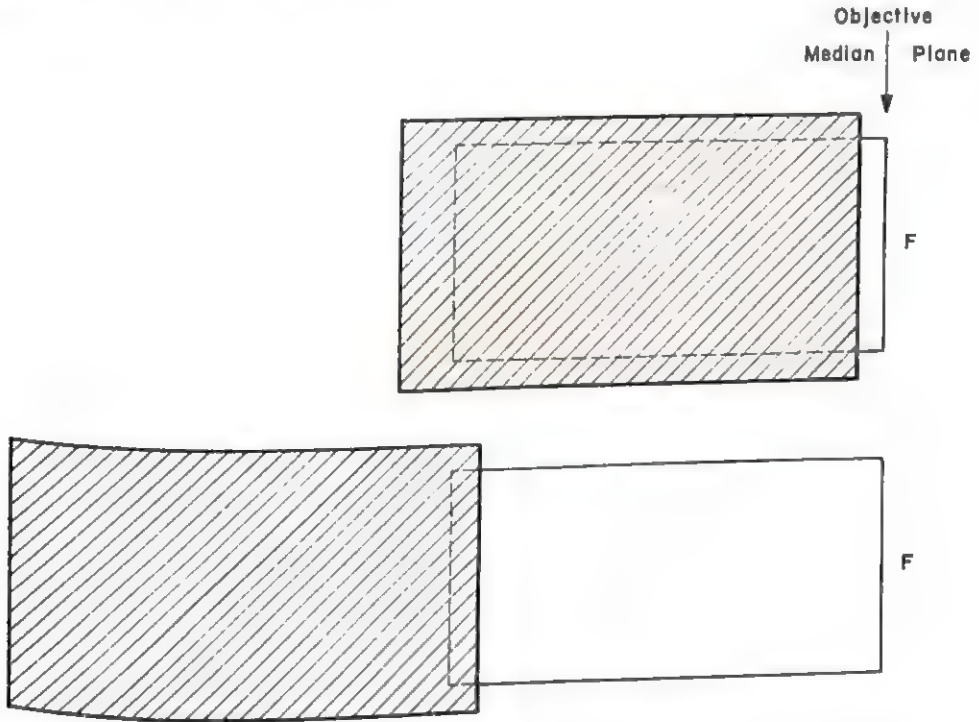


FIGURE 1
OBSERVER FIXATES EDGE (F) OF LUMINOUS STIMULUS WHILE WOODEN PANEL IS MOVED
FROM INITIAL POSITION (ON TOP) TO END POSITION (BELOW), AND THE STIMULUS
Appears TO MOVE TO THE RIGHT

D. DISCUSSION

The illusion described has relevance to the problem of retinal local signs. The common assumption is that an object placed in the objective median plane of an observer and fixated by him will be seen as straight ahead. In other words, it is usually believed that with symmetrical convergence, the fovea has the local sign of straight-aheadness.

Such a situation was created in our experiment. Eye position was kept constant throughout the experiment as *S* fixated an object placed in his objective median plane. Convergence remained symmetrical. Yet, contrary to conventional expectation, the local sign of the fovea did change. With

a vertical line as stimulus, the fixated line was seen as straight ahead; the fovea had the local sign of straight-aheadness. When, however, a luminous area to the left of the fixated line was exposed gradually, the fixated right edge of this area appeared to *S* to move to his right; the fovea changed its local sign from straight-aheadness to rightness, in spite of the fact that eye position and stimulation of the fovea remained constant.²

This observation is contrary to any theory of retinal local signs which would try to explain phenomena of localization solely in terms of the two factors of eye position and retinal area stimulated. In a recent paper (1), we have tried to explain phenomena of visual localization in terms of underlying processes of involuntary and voluntary innervation of the extraocular muscles. We have postulated that if fixation of an object can be maintained by reflex innervation of the eye-turners alone, the object will be seen straight-ahead. If, on the other hand, fixation of an object requires voluntary innervation of the levo-rotators or the dextro-rotators of the eyes, the fixated object will be seen to the left or right respectively. In other words, we suggested that the local sign of the fovea changes with the state of oculomuscular innervation. We assumed that the fovea has a local sign of straight-aheadness when the eye-turners receive reflex innervation only (*Hypothesis 1*); and we assumed that the fovea changes its local sign to leftness with voluntary innervation of the levo-rotators and to rightness with voluntary innervation of the dextro-rotators (*Hypothesis 2*).

These hypotheses can be applied easily to the phenomena described in the present paper. When a line placed in the objective median plane is the only visible object in an otherwise dark field, fixation of such a line is maintained due to the fixation reflex and due to postural reflexes which tend to keep the eyes in a position symmetrical with respect to the median sagittal plane of the body. According to our first hypothesis, under such conditions the fovea has a local sign of straight-aheadness. When, however, such a line becomes the right edge of a luminous area expanding toward the left, stimulation of receptors more and more distant from the fovea leads to an increasing reflex tendency to turn the eyes toward the peripheral stimulus. Therefore, fixation of the right edge of the stimulus can be maintained only by voluntarily counteracting the reflex tendency to turn the eyes to the left.

²For the purposes of simplicity, our discussion is limited to what happens to the local sign of the fovea. Actually, when the fovea changes its local sign, all retinal points change their local signs. When the local sign of the fovea changes from straight-aheadness to laterality, a peripheral point on the retina assumes the local sign of straight-aheadness. In terms of experience, not only the fixated edge, but the whole stimulus appears to move to the left or right.

As the levo-rotators receive increasingly strong *reflex* innervation, the dextro-rotators must receive *voluntary* innervation, the magnitude of which must increase correspondingly. According to our second hypothesis, such voluntary innervation of the dextro-rotators changes the local sign of the fovea to rightness. The fixated object is seen more and more right as the magnitude of the voluntary innervation increases. The stationary object appears to move to the right. Similarly, when the stimulus situation requires a progressive increase of voluntary innervation of the levo-rotators, the fovea changes its local sign gradually from straight-aheadness to leftness. A fixated stationary object will appear to move to the left. Such a situation was created in that condition of our experiment in which the left edge of a rectangle expanding gradually toward the right had to be fixated.

E. SUMMARY

A new illusion of apparent movement was described. A fixated (foveally projected) vertical line placed straight ahead of an observer is seen as straight ahead. If the line becomes the right edge of a rectangle by exposing gradually a luminous area to the left of the fixated line, apparent movement of the line to the right is experienced. Similarly, a vertical line appears to move to the left when it is the left edge of a luminous rectangle gradually expanding to the right. This phenomenon was discussed in relation to the concept of retinal local signs.

REFERENCES

1. BRUELL, J. H., & ALBEE, G. W. Notes toward a motor theory of visual egocentric localization. *Psychol. Rev.*, 1955, **62**, 391-400.
2. ———. Effect of asymmetrical retinal stimulation on the perception of the median plane. *Percept. & Motor Skills*, 1955 (in press).
3. ROELOFS, C. O. Optische Lokalisation. *Arch. Augenheilk.*, 1935, **109**, 395-415.

Department of Psychology
Western Reserve University
Cleveland 6, Ohio

CHANGES IN FLICKER-FUSION THRESHOLD AND IN CHOICE REACTION TIME INDUCED BY ELECTROCONVULSIVE THERAPY* †

The Psychiatric Institute, New York

CARNEY LANDIS, DONALD DILLON, AND JULIUS LEOPOLD

A. INTRODUCTION

It has been shown that some psychiatric patients who receive electroconvulsive therapy (*ECT*) show during the course of and immediately after treatment a lowered flicker-fusion threshold (*CFF*) and a decrease in the speed of reaction time (12). Berg (2) introduced a promising modification in experimental procedure for accentuating such *CFF* changes. Impressed by the individual reliability of *CFF* measures and by the wide variations of this threshold between different persons, he sought a method which would indicate in any given patient the effect of brain injury on the *CFF*. He theorized that an agent which might depress the *CFF* "might possibly disclose a latent weakness in the functional capacity of the nervous system." After investigation with several depressant drugs, he chose evipan,¹ one of the hypnotic barbiturates, which has a rapid onset and a short duration of effect. He found that an intravenous injection of a 10 per cent dose of evipan produced in about a minute after injection "an agreeable lassitude and mild euphoria" which persisted for 10 to 20 minutes. With a group of 68 non-brain injured persons he found that the mean change (decrease) in *CFF* brought about by the 10 per cent drug dosage was 0.25 F.U.² (0.16 cps \pm 0.11 cps). He concluded that a change of more than 2.0 F.U. (1.26 cps) was beyond the range of change induced by this drug dose in healthy persons. He made 77 determinations of the evipan effect on *CFF* among 20 patients during the course of *ECT* treatment. He found that evipan induced in every instance a change of less than 2.0 F.U. (1.26 cps) before

*Received in the Editorial Office on September 5, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

†This investigation was supported, in part, by research grants from the Carnegie Corporation of New York and by grant M-872 from the National Institute of Mental Health, United States Public Health Service.

¹Evipan is hexobarbital soluble and marketed in the USA under the trade name of Evipal Sodium.

²Berg varied luminance and held frequency constant. A filter unit (F.U.) was equivalent to log 0.06 ft. candles or to 0.63 cps.

ECT treatment. In 76 of 77 determinations made when the patient was tested 6 to 7 hours after an *ECT* treatment,³ the *CFF* was lowered from 2 to 4 F.U. (1.26 to 2.52 cps; mean = 1.81 cps). He also applied the evipan test to groups of patients who were suffering from head injury, poliomyelitis, CO intoxication, toxemia of pregnancy, disseminated sclerosis, and neurosyphilis. His results indicated a high degree of correlation between the extent of presumed brain injury or pathology and the course of recovery from the injury.

Åkerlund (1) utilized Berg's evipan method in a study of 239 cases of head injury, obtaining positive results with 209 patients. He also used this test with pugilists before and after boxing matches, finding good agreement between the number and the severity of blows to the head and positive evipan *CFF* changes.

Blomberg (3) studied 53 head injury patients, finding only nine who gave positive evipan *CFF* responses. He attributed the positive results obtained by Berg and Åkerlund to unconscious suggestion on the part of the experimenter.

The use of 0.4 mg nitroglycerin, sublingual absorption, as a method of inducing changes in *CFF* was introduced by Krasno and Ivy (10). They reported that nitroglycerin produced a transient drop in *CFF* in a normal healthy population and an increase in *CFF* in a group of patients suffering from cardiovascular disease.

The effects of nitroglycerin and the cold pressor test (immersion of the hand and forearm in ice water) were investigated by Landis, Clausen, and Gjesvik (11) with three normal subjects. They found that the nitroglycerin did not produce a regular drop in *CFF*, but that in 26 out of 27 instances the cold pressor test produced a rapid rise in *CFF*.

Landis and Clausen (12) studied the changes in *CFF* and reaction time with a group of nine patients who received *ECT*. They failed to find statistically significant changes in either measure during the course of, or after, the therapy. They attributed their failure to the fact that most of this small group of patients were mentally depressed, as well as not being typical of the usual range of psychiatric patients.

Since evipan, nitroglycerin, and the cold pressor response all involve a series of accentuated physiological changes, they may be spoken of as physiological loads.

³In 7 instances, the test was not made for 24 hours after treatment. The change in these seven was of the same average amount as the change which occurred 6 to 7 hours after treatment.

B. PURPOSE

In the light of these considerations it seemed worth while to investigate (a) the effect of *ECT* on *CFF* and reaction time, and (b) to utilize the evipan, nitroglycerin, and cold pressor methods to see whether these physiological "loads" might accentuate or otherwise clarify the changes in *CFF* or reaction time measures which might be induced by *ECT*.

C. APPARATUS

1. *Flicker-Fusion Threshold*

The same apparatus was employed which was described by Landis, Clausen, and Gjesvik (11). Essentially this was an electronically controlled device which presented an intermittently illuminated test patch. The control circuits were designed so that the rate of intermittence could be started at 70 cps or any multiple of 5 between 70 and 25 cps. From a preselected starting frequency an oscillator automatically decreased the rate of intermittence one cycle per second, each second of elapsed time. The preselected starting frequency was one which had been found to be high enough so that the test patch was first seen as a steady light by the particular observer. He was instructed to watch the lighted patch closely and as soon as he saw the light start to flicker (a pronounced flickering patch was demonstrated), he should press a switch which would extinguish the light. Pressing this button locked the frequency which was being delivered by the oscillator. This locked frequency was electronically "scanned" and the result displayed on a bank of digital counters and printed as a three-digit number (cps to a tenth of a cycle) on a paper tape by an R-3 Berkeley Digital Recorder. As soon as the number was printed the apparatus cleared and reset itself to the preselected starting frequency and was ready to be started again by the experimenter. The patient was seated 18 inches from the test patch. His head was held in position by a chin rest. The test patch was a round, white-flashed opal glass surrounded by black felt arranged so that only the black surround was visible during observation. The test patch subtended 2° of visual angle at the cornea and had a luminance of 9 ml when seen as a steady light. The light-dark ratio was 0.50 and the light pulse was of a square wave form. In any single session 20 successive "descending" thresholds were obtained, the score being the mean of these 20 expressed in cycles per second. The accuracy of the instrument was 0.01 cps.

2. Reaction Time-Tapping Equipment

A newly designed apparatus was built for these measurements.⁴ A schematic block diagram of two of the circuits of this equipment is shown in Figures 1A and 1B while Figure 2 is a sketch of the equipment in operation. The stimulus situation was the same as that described by Landis and Clausen (12). The instructions to the patient were as follows:

Place your index finger on this Home button. You will hear a buzzer sound which means "Ready," then in a couple of seconds after the buzzer either the red or the green light will flash. Move your finger just as rapidly as possible to the target plate which is under the light which has just flashed. Hold your finger on the target a couple of seconds and then move your finger back to the Home button.

The time elapsing between the flash of the red or the green light and the instant the finger lost contact with the starting button was measured in milliseconds and displayed on a bank of Berkeley Digital Counters. This measure of elapsed time we termed decision time or *D*-time. The time elapsing between lifting the finger from the Home button and touching the target contact was measured in milliseconds and displayed on an additional bank of digital counters. This measure we have termed ballistic time or *B*-time. Twenty *D*-time and *B*-time determinations were made at each test session.

The patient was then told to place his finger on the right target contact and when the green light flashed to move his finger as rapidly as possible back and forth between the red and green target contacts. The apparatus was adjusted to count the number of taps made in a preselected three-second interval. This number was displayed on the digital counters and was termed "Number of taps." Twenty trials were taken at each test session.

D. SUBJECTS

This equipment was installed in two rooms in the Reception Building of the Creedmoor State Hospital, Queens Village, New York.⁵ All newly admitted patients whose condition permitted were given the *CFF*, *D*-time, *B*-time, and Number of taps tests as soon as possible after entering the hospital. Valid initial measurements were obtained from 110 patients. The measures obtained at the first test session we have called *Pre-ECT* scores.

⁴This equipment was designed and built by Dr. Fred H. Ireland.

⁵We are indebted to Dr. Harry LaBurt, Director of Creedmoor State Hospital, and to members of his staff for their cooperation and assistance in the collection of this data. We are further indebted to Mr. Alexander Van West and to Mr. Bert Kaplan for assistance in the collection of this data.

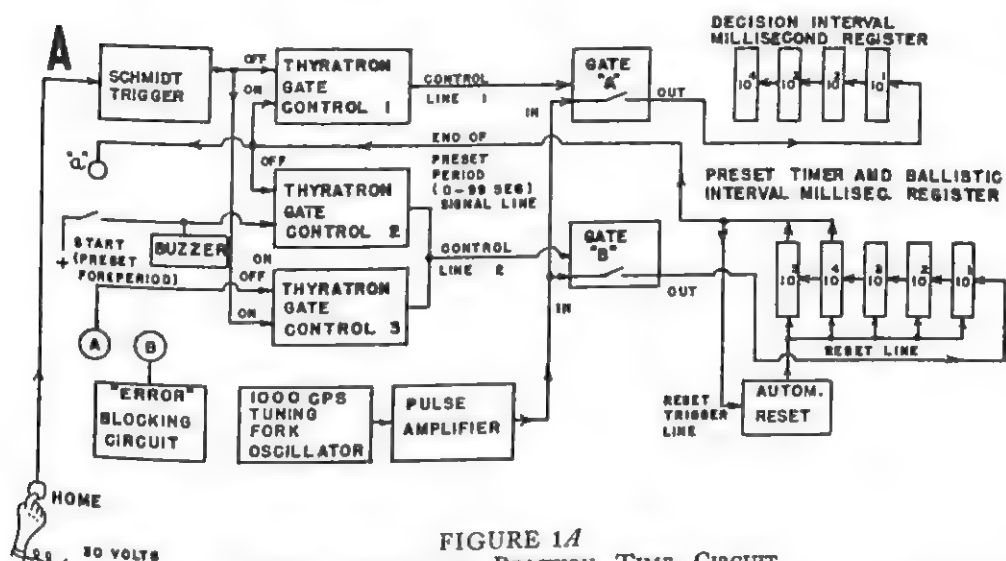


FIGURE 1A

BLOCK DIAGRAM OF REACTION TIME CIRCUIT.

This is set for a correct response. Preset counters may be set for 1-99 sec delay period. To give ready signal and start cycle, close Start switch. This sounds buzzer, throws Thyatron gate control No. 2 "on," opens gate "B" and permits Preset interval timer to register passage of the 1,000 cps impulses. At end of preset time, signal goes over "End of preset period signal line" which (1) shuts off Thyatron gate No. 2 and closes gate "B"; (2) illuminates stimulus light "a"; (3) triggers automatic reset circuit which resets bottom counters to zero; (4) turns Thyatron gate control No. 1 on which opens gate "A" thus allowing Decision time interval register to count Decision time. Subject lifts finger from "Home" button. This activates Schmidt trigger which emits signal; (1) which shuts off Thyatron gate control No. 1, which opens gate "A" and stops Decision interval register; (2) turns on Thyatron gate No. 3, closes gate "B" and enables Ballistic interval register to count the 1 kc pulses. Subject touches correct target contact "A" under "a" stimulus light which shuts off Thyatron gate control No. 3 so opening gate "B" and stopping Ballistic interval counter. If subject reacts to incorrect target contact "B," error blocking circuit will be energized indicating an error in the decision response.

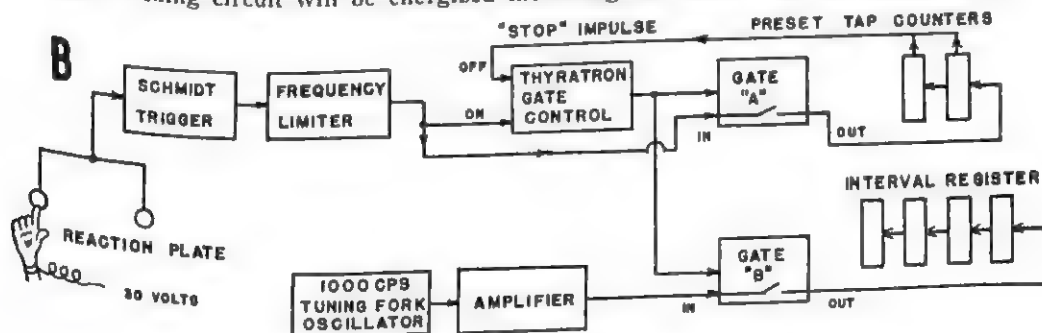


FIGURE 1B

BLOCK DIAGRAM OF TAPPING CIRCUIT

First tap activates Schmidt trigger sending an impulse through Frequency limiter to Thyatron gate control which opens Gates A and B. The interval timer immediately begins to count the passage of 1,000 cps impulses. Succeeding taps occurring within frequency limiter response band are sent through Gate A and are registered by Preset tap counters. When tap counters have registered the number of taps for which they were preset, they send "stop" impulse to "Thyatron gate control" which in turn closes Gates A and B. Preset tap counters show number of intervals counted; interval register shows total time in milliseconds needed to accomplish the predetermined number of taps.

The hospital course of each patient was followed. All patients who had not had *ECT* or insulin coma treatment during the previous year, who were cooperative, and who were placed on *ECT* therapy were tested again after receiving between 6 and 12 (usually 10) treatments. This second test session

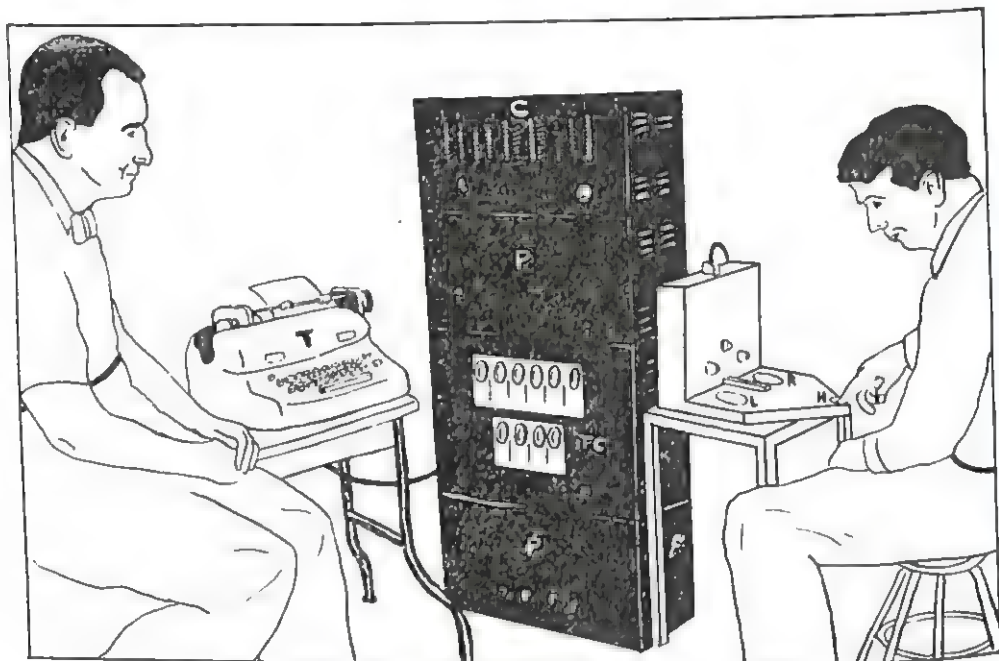


FIGURE 2
REACTION-TIME-TAPPING APPARATUS

H = home contact; *L* = left or red target contact plate; *R* = right or green target contact plate; *C* = digital counters; *P* = power supply; *TC* = typewriter control; *T* = electric typewriter which recorded the score for each performance.

has been termed the "During" session. At the end of the *ECT* treatments (usually 20) each patient was again tested; this third session we termed "Final." If the patient continued to be available in the hospital and if his status was uncomplicated by other forms of active therapy he was retested at intervals of 2, 3, 4, and 6 weeks after the termination of *ECT* treatments. These have been termed Post-2 wk, Post-3 wk, Post-4 wk, and Post-6 wk, respectively.

Four varieties of test sessions were employed; (a) *Evipal, regular*:⁶ The tests were administered in the morning after breakfast before any medication had been given the patient and before the patient had been allowed to smoke.

⁶The dosage and drug *Evipal* which we used were the same as the one-tenth dose *evipan* of Berg.

After the battery of tests had been completed, the ward physician administered intravenously a one-tenth dose of Evipal, the amount being determined by the dosage system of Specht (16). Twenty *CFF* determinations were made during a period of five mins after the injection. Immediately thereafter, the reaction-time and tapping tests were given. The three tests took 15 to 20 mins to complete, which is within the period of the effectiveness of the injected Evipal. The During and Final test session measurements were made approximately 24 hours after the most recent *ECT* treatment. (b) *Evipal, special*: The procedure here was the same as that for the Evipal regular, except that the During and Final test sessions were completed 5 to 6 hours after the morning *ECT* had been administered. (c) *Nitroglycerin*: This procedure was the same as that for Evipal, regular, except that the patient was given a 0.43 mg (1/100th grain) tablet of nitroglycerin which he allowed to dissolve under his tongue in place of the injections of Evipal. (d) *Cold pressor*: The procedure here followed that used in *Evipal, special*, namely tests were done on the same day as *ECT*. In place of Evipal injections the hand of the patient was immersed in ice water during both the *CFF* and the reaction time tests. The period during which the tapping test was administered served as a buffer since it was given between the *CFF* and reaction time tests without having the hand in ice water.

As in every experimental study employing psychiatric patients, there were numerous occasions when measurements could not be secured so that the identity and the number of patients comprising the relevant groups usually differed somewhat at each test period. One may present an analysis of the complete group of data secured at each test period disregarding the fact that the number of patients and the identity of the patients in a particular group are various. Or one can select for presentation the "solid" data which include only the measurements obtained from patients who participated in all tests at each test period, so rejecting all "fragmentary" data. Landis and Clausen (12) analyzed similar data which they had secured, and found that the results obtained when they used the solid data method were usually comparable to those obtained when complete plus fragmentary data were used. In light of this we have utilized the data from all patients from whom we obtained test scores. When statistical comparisons required solid data, we have restricted the group and have noted the restriction.

In Table 1 is given the relevant data as to diagnosis, sex, age, and outcome for the four experimental load groups. We tested the data which we obtained to see if sex, age, diagnosis, or outcome influenced the results in any consistent fashion. The correlation between age and *CFF* was -0.39

TABLE 1
DIAGNOSIS, SEX, AGE, AND OUTCOME AT FINAL ECT SESSION ACCORDING TO LOAD UTILIZED

Group	Diagnosis				Sex		Mean age	Age range	Outcome at final test session			
	Dementia Præcox	Involutional Psychosis	Manic Depressive		Male	Female			Unimproved	Slight improvement	Improved	Total
<i>Load:</i>												
Evipal regular	24	4	1		16	13	37.2	21-59	6	13	10	29
Evipal special	8	3	0		10	1	38.1	19-54	6	4	1	11
Nitroglycerin	14	6	1		15	6	40.5	22-62	9	8	4	21
Cold pressor	11	0	1		8	4	30.6	19-46	6	4	2	12
Total	57	13	3		49	24	37.2	19-62	27	29	17	73
<i>Sex:</i>												
Male	39	9	1		49	—	35.6	19-62	24	15	10	—
Female	18	4	2		—	24	40.5	27-59	3	14	7	—
<i>Age:</i>												
Mean age	33.5	50.8	49.7		—	—	—	—	32.2	38.7	42.5	—
Age range	19-59	40-62	44-59		—	—	—	—	19-50	21-62	26-59	—
<i>Outcome:</i>												
Unimproved	26	1	0		24	3	—	—	—	—	—	—
Slight improvement	19	8	2		15	14	—	—	—	—	—	—
Improved	12	4	1		10	7	—	—	—	—	—	—
Total	57	13	3		49	24	—	—	—	—	—	—

($p < .01$) and between age and changes produced by the loads was -0.19 ($p > .05$). There seemed to be some tendency (though not significant) for the *CFF* measures obtained at the Pre-*ECT* session to be associated with outcome, but when age was held constant this tendency became less rather than greater.

E. RESULTS

1. *The Effect of Electroconvulsive Therapy*

Each patient was tested twice at all test sessions, first without a physiological load and then again (during the following 90 mins) with load, namely, after absorption of 0.40 mg nitroglycerin, intravenous injection of one-tenth dose of Evipal, or while holding one hand immersed in ice water (cold pressor test). The scores obtained without load at the During test session (after 6-12 *ECT* treatments) and at the Final session (after the final *ECT* treatment) represent the effect of *ECT* on these measures. In Table 2 the group means and standard deviations, together with the average *SD* of the individuals composing each group, obtained at each test session (without load) are given for each of the test scores. The mean scores for *CFF*, *D*-time, and choice reaction-time (*D*- + *B*-time) were changed significantly ($p < 0.01$) when the During and Final scores were compared to the Pre-*ECT* scores. The *SD* measures for these tests were also markedly increased for the During and Final sessions.

The mean and *SD* of the scores for *B*-time and Number of taps were not altered significantly at the During and Final test sessions. Further analysis showed that the *B*-time and Number of taps scores were not affected significantly by any of the loads employed. It was found that in certain of the load comparisons the summing of the *D*-time and *B*-time scores (*D*- + *B*-time), which gives a measure that traditionally has been called choice reaction time, gave an increased statistical significance to certain comparisons over that given by *D*-time alone. Hence, we have dropped from this report any further consideration of the *D*-time, *B*-time, and Number of taps scores and have dealt only with the *CFF* and choice reaction time (*CRT*) scores.

Since the group means and *SD*s for *CFF* and *CRT* were changed significantly by *ECT*, we were interested in the number of patients who gave rise to these changes. Of the 58 patients from whom we obtained *CFF* measures at the During session, 76 per cent showed a lower score. Of the 59 patients from whom *CRT* measures were secured, 66 per cent had a slower score. Of the 43 patients (*CFF*) and 45 patients (*CRT*) at the Final session, these percentages were 79 and 62, respectively. Using complete

data and the method of fiducial limits for determining significance (14), it was shown that the change (loss) at Final compared to Pre-ECT session was significant at the .05 level of probability for 71 per cent of the individuals for their *CFF* score and for 54 per cent for their *CRT* score. Using solid data and critical ratio method (14) for determining significance, the percentages were 74 and 47, respectively.

TABLE 2
NUMBER OF PATIENTS, MEAN, GROUP STANDARD DEVIATIONS, AND AVERAGE *SD* FOR THE INDIVIDUALS FOR EACH TEST AT EACH TEST SESSION WHEN NO LOAD HAD BEEN UTILIZED. * ($p < 0.01$)

	Pre-ECT	During	Final	Session			
				Post-2 wk	Post-3 wk	Post-4 wk	Post-6 wk
Flicker-fusion score in cps							
<i>N</i>	58	58	48	53	53	52	43
Mean	41.26	39.52*	39.06*	42.32	42.03	42.66	42.37
Group <i>SD</i>	4.17	4.73	4.73	3.95	3.98	3.52	3.85
Average <i>SD</i>	.88	.99	1.27	.81	.78	.78	.76
Decision time in milliseconds							
<i>N</i>	59	60	55	56	55	54	45
Mean	379.9	459.8*	468.4*	349.3	338.2	327.7	317.5
Group <i>SD</i>	120.0	204.3	203.0	81.6	77.3	69.0	59.4
Average <i>SD</i>	81.2	122.4	134.7	64.8	62.2	51.6	48.8
Ballistic time in milliseconds							
<i>N</i>	59	60	55	56	55	54	45
Mean	257.1	259.5	248.7	211.2	221.4	217.8	214.2
Group <i>SD</i>	97.7	106.0	106.1	82.2	83.3	86.6	75.2
Average <i>SD</i>	44.5	45.5	43.3	36.9	30.4	28.7	26.4
Choice reaction time (D- + B-time) in milliseconds							
<i>N</i>	59	60	55	56	55	54	45
Mean	637.2	720.2*	717.2*	563.0	554.3	536.3	531.7
Group <i>SD</i>	203.0	282.0	279.0	151.0	146.0	161.0	118.0
Average <i>SD</i>	79.5	132.5	138.0	77.3	70.8	60.1	56.3
Number of taps in three seconds							
<i>N</i>	59	60	54	56	55	54	45
Mean	13.8	15.2	15.6	16.5	16.6	16.6	17.0
Group <i>SD</i>	3.8	3.5	3.6	3.3	3.4	2.9	2.6
Average <i>SD</i>	1.1	1.0	1.2	1.2	.9	.9	1.0

In brief, one may say that *ECT* brought about an increased variability and a significantly lowered flicker-fusion threshold and choice reaction-time in the group scores of psychiatric patients. In terms of individual patients a lower *CFF* score (statistically significant) was shown by about seven out of every 10 patients and a prolonged *CRT* score (statistically significant) was shown in about half of the patients. Two weeks after *ECT* had been discontinued neither *CFF* nor *CRT* scores showed significant differences (residual effects) in comparison to scores obtained before *ECT*.

2. *The Load Effect of Evipal*

It will be recalled that the load effect of Evipal (regular) referred to tests made 24 hours after the *ECT* treatment while Evipal (special) referred to tests made 5 to 6 hours after the *ECT* treatment. Since some of the patients appeared somewhat confused or dazed 5 to 6 hours after *ECT*, it was thought that these two methods might yield different results at the During and Final sessions. Examination of these data showed that there were no statistically significant differences in the group scores between the regular and special procedures. Hence, we have combined these data and analyzed them as Evipal load irrespective of whether the test was given 5-6 hours or 24 hours after *ECT*.

Table 3 gives the mean *CFF* score in cps and mean *CRT* score in ms for each test session without load and with load, together with the number of patients whose scores showed a gain or a loss in either score after Evipal had been injected. The increase in the group mean *CFF* after Evipal injection at the Pre-*ECT* session amounted to 0.7 cps. This plus the fact that 22 of 36 patients showed an increased score at this session was the only comparison of particular interest but neither of these changes were statistically significant. The Evipal produced a somewhat speedier *CRT* group mean at the first three test sessions but the change was not statistically significant at any session either for the group or for the number of individuals.

Berg (2) reported that in 76 out of 77 instances where *CFF* was measured after *ECT* the *CFF* was 2 F.U. (1.26 cps) lower after one-tenth dose of evipal had been administered than when no evipal had been given. This method he believed accentuated the changes induced by *ECT* since the lowered score did not occur before *ECT* had been given and the differential did not last more than six days after the final *ECT* treatment.

Had our Evipal-*CFF* findings been in accord with Berg's, the number who gained and who lost at the During and Final sessions (Table 3) should have been something like 1 and 33 and 1 and 29, rather than 18 and 16 and 16 and 14 which we did obtain.

Actually there were 172 individual measures of *CFF* after one-tenth dose of Evipal had been given and where *ECT* had not been administered during the previous 14 days. These may be compared to 65 instances when Evipal was administered within 24 hours after an *ECT* treatment. Of the 172 instances there were 76 (44 per cent) that showed some loss in *CFF* score and of these 76, 36 (21 per cent) had lost more than one *SD* of their score at the "No load" test. Of the 172 instances, 126 showed losses of

TABLE 3

THE MEAN GROUP SCORES FOR CFF AND CRT WITHOUT LOAD AND WITH THE LOAD EFFECT OF 1/10 DOSE EVIPAL, OF 0.40 MG NITROGLYCERIN, OR OF THE COLD PRESSOR REACTION AT EACH SESSION TOGETHER WITH THE NUMBER OF INDIVIDUALS WHO SHOWED GAINS OR LOSSES WHEN THEIR SCORE WITH LOAD WAS COMPARED TO THEIR SCORE WITHOUT LOAD

Load	Session											
	Pre-ECT			During			Final			Post-2 wk		
	No	Load	No	No	Load	No	No	Load	No	No	Load	No
<i>Evipal</i>												
Mean CFF	41.5	42.2	40.2	40.5	39.3	39.8	42.8	42.7	42.3	42.4	42.8	42.5
No. who gained		22	18	18		16	17	17	18	18	14	12
No. who lost		14	16	16		14	15	15	15	15	17	15
Mean CRT	648	614	697	686	745	695	578	587	558	553	533	543
No. who gained		22	18	18		20	18	18	20	20	10	10
No. who lost		12	18	18		13	18	18	13	13	24	24
<i>Nitroglycerin</i>												
Mean CFF	40.4	41.5	38.1	38.6	39.4	39.4	41.7	41.7	41.2	41.3	42.1	42.5
No. who gained		11	10	10		4	6	6	6	6	10	10
No. who lost		4	9	9		7	6	6	8	8	6	8
Mean CRT	720	700	820	812	715	762	535	565	579	585	588	592
No. who gained		16	7	7		6	4	4	7	7	8	8
No. who lost		6	15	15		14	13	13	9	9	10	10
<i>Cold Pressor</i>												
Mean CFF	42.3	42.7	38.9	39.3	38.7	38.4	39.5	39.6	39.1	39.0	—	—
No. who gained		7	8	8		1	3	3	2	2	—	—
No. who lost		4	4	4		7	4	4	3	3	—	—
Mean CRT	633	618	669	684	591	620	621	735	518	585	—	—
No. who gained		5	6	6		4	1	1	1	1	—	—
No. who lost		7	6	6		4	5	5	4	4	—	—

more than 1.26 cps, which figure Berg had considered beyond the range of variability in non-brain injured persons. That is, a blind diagnosis on the basis of loss in *CFF* after injection of Evipal (uncomplicated by *ECT*) would have resulted in a 15 per cent incidence of false diagnosis of brain injury.

Of the 65 determinations of *CFF* made within 24 hours of an *ECT* treatment, there were 34 gains, 30 decreases, and one no change in scores. Of the 30 decreases, 16 (53 per cent) were greater than one *SD* of the mean score of the "No load" test. Of the 65 determinations only 14 (22 per cent) were decreases of 1.26 cps or more. That is, 22 per cent of the *CFF* measures after *ECT* met Berg's criterion, while 78 per cent failed to do so.

There were 171 instances where *CRT* measures were made after Evipal injection and where *ECT* had not been administered during the preceding 14 days. Of these, 48 per cent showed an increased speed of response (fewer milliseconds of elapsed time) after Evipal injection, and 49 per cent a slower response, while 3 per cent were unchanged. In the 70 instances where the Evipal comparisons were made within 24 hours after *ECT*, 54 per cent were faster, 44 per cent were slower, and 2 per cent were unchanged. Evidently, Evipal had no differential effect on *CRT* no matter whether *ECT* intervened or not.

3. *The Load Effect of the Cold Pressor Test*

Hines and Brown (6) first demonstrated that holding the hand in ice water resulted in a quick rise in blood pressure. Steinhaus and Kelso (17) reported that a cold hip bath raised the *CFF* in most healthy persons. Landis, Clausen, and Gjesvik (11), using three healthy observers, found in 26 of 27 comparisons that the *CFF* was raised by the cold pressor test.

The results of the use of this test in the present study are summarized in Table 3. The *CFF* and *CRT* group means at every test session were very nearly the same (considering the size of the *SD*) during the cold pressor test. The number of individuals who gained or lost on *CFF* was changed from 7-4 to 1-7 which was not significant (χ^2 test), when Pre-*ECT* was compared to Final (*ECT*) session. Taking the 23 instances where the cold pressor test was administered uncomplicated by *ECT*, there were 12 increases and 11 decreases. Of the 20 tests where comparisons were made at the During and Final sessions, 9 increased and 11 decreased.

For the 23 *CRT* comparisons not complicated by *ECT*, 7 (30 per cent) gained and 16 (70 per cent) lost in speed, while for the 20 comparisons

obtained at the During and Final (*ECT*) sessions there were 10 who lost and 10 who gained. The change in distribution of gainers and losers was not significant.

In general the cold pressor test did not change the *CFF*. It tended after the Pre-*ECT* session to prolong *CRT* time, but the tendency was not statistically significant. The use of the cold pressor reaction as a device to accentuate the changes in *CFF* and *CRT* produced by *ECT* was not successful.

4. *The Load Effect of Nitroglycerin*

Krasno and Ivy (10) reported that among 206 patients without evidence of cardiovascular disease all gave a decrease in *CFF* within four to five minutes after a 0.4 mg nitroglycerin tablet had been dissolved under the tongue. Among 216 patients with various cardiovascular conditions all showed an increased *CFF* after sublingual absorption of 0.4 mg nitroglycerin. Subsequent investigators (4, 5, 7, 8, 9, 11, 15) have reported that in persons free from cardiovascular disease from 21 per cent to 94 per cent of the groups investigated gave a decrease in *CFF* in response to the administration of nitroglycerin, while among groups of patients suffering from known cardiovascular disease from 48 to 97 per cent of them showed an increase in *CFF* following the administration of nitroglycerin.

All patients in the present group were free from any known evidence of cardiovascular disease. Table 3 presents the load effect of nitroglycerin at each test session for both *CFF* and *CRT*. It is evident that nitroglycerin did not affect the mean *CFF* scores either when *ECT* complicated or did not complicate the comparisons. There was some shifting about of the number of individuals who gained or lost at particular test sessions, but none of these shifts was statistically significant. Among this group of 19 patients there were 78 occasions where *CFF* comparisons, with and without nitroglycerin, could be made where *ECT* had not occurred during the previous two weeks. Of these 78 comparisons, 36 (46 per cent) were increases, 32 (41 per cent) decreases, and 10 did not change. There were 33 instances where the test was made within 24 hours after *ECT*; of these 14 (42 per cent) showed a gain, 16 (48 per cent) a loss, and three showed no change. Evidently whether nitroglycerin raised or lowered *CFF*, either with or without the complication of *ECT*, was a matter of chance.

Nitroglycerin as a load produced certain changes in choice reaction time scores which were unexpected. The change in the *CRT* group mean score before and after absorption of nitroglycerin was not significant at any test session. The *SDs* of the means for the measures with or without load for

the Pre-ECT, During, and Final test sessions ranged from 230 to 300 ms. The average of the individual *SDs* for these sessions ranges from 122 to 180 ms. The group mean *SDs* and average individual *SDs* for all four Post-sessions were less than half as great as for the first three sessions. Evidently, the *CRT* scores were subject to practice and had, comparatively speaking, a large variability. However, at the Pre-ECT test session nitroglycerin brought about a faster reaction time in 16 persons, and a slower time in six persons. This distribution of the number of persons who gained or lost was reversed at all the following test sessions and in each instance the comparison with the Pre-ECT was, by the χ^2 test, statistically significant at the .05 level of probability. (The Pre-ECT comparisons with Post-3 wk and Post-4 wk were not statistically significant.)

The *CRT* measures for individual patients were next examined to see whether the changes in *CRT* scores following nitroglycerin were significant or not for each particular patient. At the Pre-ECT session nitroglycerin was followed by an increased speed of reaction-time greater than the *SD* of the measures without the load in three of the 22 patients. At the During and Final sessions where 22 and 20 persons, respectively, were involved, three were slowed down more than one *SD* by the nitroglycerin. At the Post-2 wk session (17 patients), one gained and five lost more than one *SD* under the load. If this nitroglycerin effect on *CRT* is referred to group means, or to the number of persons whose scores were changed more than one *SD* then it would be concluded that few significant changes in reaction time could be attributed to nitroglycerin. But despite the lack of this more complete evidence the fact remains that the Pre-ECT distribution of individuals who gained or lost during the period of nitroglycerin effect was reversed at the During and Final sessions, and the reversals lasted for at least six weeks after the ECT treatments. We have, at present, no explanation to offer.

F. DISCUSSION

Two points deserve comment; first, the effect of ECT on the flicker-fusion threshold and on reaction time, and, second, the failure of the load procedure to produce an accentuation of test scores. In a previous report from this laboratory (12) we found that ECT did not alter *CFF* or reaction time in a significant fashion. This was attributed to the fact that the sample of patients was small and that most of the group were suffering from depressions. The present investigation indicates that seven out of 10 patients showed a significant decrease in *CFF* both after six to 10 ECT treatments and after 20 treatments. Based on this sample of 63 unselected psychiatric

not a procedure which can be done in a casual fashion by a naïve experimenter. Getting thresholds is deceptively easy. As a usual thing an observer will give 10 or more successive measures which vary but little among themselves, so adding to the deception. Only when one realizes the number of demonstrated determinants of this threshold, the fact of the wide variation in threshold between individual observers, the fact that even the most experienced and reliable observer will occasionally provide a series of measures quite outside of his ordinary performance for which no basis of explanation can be found; and finally the ease with which suggestion, expectation, changes in attitude and the like, may alter any series of measurements; only then can one properly evaluate a change in the level of the flicker-fusion threshold which may occur.

The reaction-time experiment is equally deceptive, particularly because of its simplicity. After a get-ready signal a light flashes and one moves one's finger as rapidly as possible over a distance of a few inches to a spot beneath the light. This seems a very simple response. But as the present experiment showed (together with some hundreds of previously published studies) there was a very large inter- and intra-individual variability in the speed of the movement.

Because of interindividual variability of both *CFF* and *CRT*, it would be highly desirable to be able to utilize a method like that which Berg, or which Krasno and Ivy, proposed, if it were available and dependable. If each person could "serve as his own control" we would be in a better position to evaluate experimental findings. There is no way at present of being certain whether a patient who responds with a *CFF* much lower than that expected under certain standardized conditions has suffered from head injury or whether he has always had a low *CFF*. When we have *CFF* and/or *CRT* measures before, during, and after *ECT* or psychosurgery, it seems safe to attribute the changes in scores to brain injury. But when we have only "after" measures, deviation from a group mean score does not form a sound basis for comparison. The search for a satisfactory reference point will have to be continued.

G. SUMMARY

Critical flicker-fusion threshold and choice reaction time measurements were obtained from a group of psychiatric hospital patients before electroconvulsive therapy, during the course of the treatments, within 24 hours after the last treatment, and at intervals of two, three, four, and six weeks after the last treatment. At each test session the patient was tested twice, once without load, and once with load. The physiological loads employed were

(a) one-tenth dose of Evipal, (b) absorption of 0.40 mg tablet of nitroglycerin, or (c) holding the hand submerged in ice water (cold pressor test) during the test performance.

Flicker-fusion and choice reaction time group mean measures were decreased significantly ($p < 0.01$) during the course of and at the end of the series of electroconvulsive treatments. The decreases in means were no longer apparent two weeks after the *ECT* treatments were finished. The decrease was shown at a significant level by 74 per cent of the patients for the *CFF* score and by 47 per cent of the patients for the *CRT* score.

The intravenous injection of one-tenth dose of Evipal did not accentuate the changes in scores brought about by *ECT*. The cold pressor test failed to change either the *CFF* or *CRT* scores in any significant fashion. The absorption of nitroglycerin did not systematically affect the *CFF*, but it did produce a change (probably significant) in *CRT* during and after *ECT* treatments. No explanation for this particular change was advanced.

REFERENCES

1. ÅKERLUND, E. A study in acute head injuries examined with flicker fusion determined under the influence of evipan. *Acta Chir. Scan.*, 1953, Suppl. 174, 1-84.
2. BERG, O. A study of the effect of evipan on the flicker fusion intensity in brain injuries. *Acta Psychiat. et Neurol.*, 1949, Suppl. 58, 1-116.
3. BLOMBERG, L.-H. Clinical studies of nystagmus produced by evipan. *Acta Psychiat. et Neurol.*, 1955, Suppl. 97, 1-88.
4. FOX, E. L., & WRIGHT, I. S. Nitroglycerin-flicker test. *Circulation*, 1951, 4, 350-358.
5. GOLDEN, L. H., & CORDASCO, E. M. Flicker photometry and electrocardiography: A correlation. *N.Y.S.J. Med.*, 1952, 52, 2157-2158.
6. HINES, E. A., JR., & BROWN, G. E. A standard stimulus for measuring vasomotor reactions: Its application in the study of hypertension. *Proc. Staff Meet., Mayo Clin.*, 1932, 7, 332-335.
7. HOPKINS, R. W. Studies in the Arctic with the flicker-fusion nitroglycerin test. *New England J. Med.*, 1952, 246, 401-403.
8. JOHNSON, R. P. The nitroglycerin-flicker fusion test. *U.S. Armed Forces Med. J.*, 1953, 4, 1167-1173.
9. KEYS, A., & SIMONSON, E. The flicker fusion frequency and its response to nitroglycerin in normal subjects and in patients with cardiovascular disease. *Circulation*, 1952, 5, 215-224.
10. KRASNO, L. R., & IVY, A. C. The response of the flicker fusion threshold to nitroglycerin and its potential value in the diagnosis, prognosis, and therapy of subclinical cardiovascular disease. *Circulation*, 1950, 1, 1267-1276.
11. LANDIS, C., CLAUSEN, J., & GJESVIK, A. Effect of varying stimulus conditions on the flicker-fusion threshold when nitroglycerin or the cold pressor test was employed. *J. Appl. Physiol.*, 1955, 7, 513-518.
12. LANDIS, C., & CLAUSEN, J. Changes in sensory and motor performances induced by active psychiatric treatment. *J. of Psychol.*, 1955, 40, 275-305.

13. LANDIS, C., & HAMWI, V. The effect of certain physiological determinants on the flicker-fusion threshold. *J. Appl. Physiol.*, 1954, **6**, 566-572.
14. McNEMAR, Q. *Psychological Statistics*. New York: Wiley, 1949.
15. RUSSEK, H. I., ANDERSON, W. H., & DOERNER, A. A. Significance of the nitro-glycerin-flicker fusion response in normal subjects and patients with cardiovascular disease. *Circulation*, 1951, **4**, 359-365.
16. SPECHT, K. Rausch-, Kurz- und Einleitungsnarkose mit Evipan-Natrium. *Zentralbl. Chir.*, 1933, **60**, 242-247.
17. STEINHAUS, A. H., & KELSO, A. Improvement of visual and other functions by cold hip baths. *War Med.*, 1943, **4**, 610-617.

The Psychiatric Institute
722 West 168th Street
New York City 32

LYSERGIC ACID DIETHYLAMIDE (LSD-25): XVII. TOLERANCE DEVELOPMENT AND ITS RELATIONSHIP TO A THEORY OF PSYCHOSIS*¹

The Biological Laboratory, Cold Spring Harbor

H. A. ABRAMSON, M. E. JARVIK, M. H. GORIN, AND M. W. HIRSCH

A. INTRODUCTION

The first observations recognized as the development of tolerance to lysergic acid diethylamide (LSD-25) are apparently those of Isbell *et al.* (6) who studied drug addicts at the NIMH Addiction Research Center. From these data it can be said that LSD-25 is its own best antidote.

The purpose of the present experiment was to determine how lysergic acid diethylamide antidotes itself, that is, evokes tolerance under two conditions: (a) high doses repeated daily and, (b) low doses increased on successive daily administration. Another purpose was to determine whether administration of d-l-brom lysergic acid diethylamide (BOL-148), an LSD-25 derivative, would evoke tolerance to LSD-25 given the following day. We have found a marked diminution of response to successive administrations of LSD-25 as measured by direct observation and by our questionnaire (2). BOL-148, however, did not noticeably inhibit the usual LSD-25 response.

Savage (8) studied the therapeutic effects of LSD-25 on depressed patients. Starting them "on an oral dose of 20 micrograms, which was increased daily to a point where a definite psychophysiological effect could be observed," he then gave that dose daily for one month and drew the following conclusion: "Improvement obtained during the course of LSD therapy was not greater than that obtained without its use in comparable cases." Our findings in the present experiment suggest that the anticipated therapeutic effects of LSD-25 did not appear due to the development of tolerance.

Lovell *et al.* (7), in reporting no side effects from LSD-25 given in increasing doses for 9-11 days, apparently were not aware of the "autodoting" effects of the drug. This phenomena can also be found by a careful study of the work of Frederking (4). Hoch, Cattell, and Pennes (5) found that

*Received in the Editorial Office on September 7, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹From The Biological Laboratory, Cold Spring Harbor Long Island, New York. This investigation was aided in part by a grant from the Josiah Macy, Jr. Foundation, New York City.

some patients who received LSD-25 responded inconsistently from experiment to experiment, while others gave similar responses. The variant results may be related to the time interval between experiments and development of tolerance.

B. METHOD

1. Subjects

Three subjects were used in this experiment. All had served as subjects at least six times and as many as 50 times before the present experiment. Subject *A* was a 26-year-old female weighing 145 pounds. Subject *B* was a 38-year-old male whose weight was 170 pounds. Subject *C* was a 39-year-old female, weight 132 pounds. All subjects were in good health, and all were college graduates.

2. Experiments

Subject *A* received 100 micrograms of LSD-25 at 10:00 A.M. on six consecutive days. Five days later she received 100 micrograms of LSD-25 at 9:25 A.M.

Subject *B* received 100 micrograms of LSD-25 at 7:00 P.M. on three consecutive days. In a second series of experiments he received five micrograms of LSD-25 at 9:30 A.M. and at 8:00 P.M. of the first day, 10 micrograms at 9:45 A.M. of the second day, and 25 micrograms at 9:45 A.M. and at 8:30 P.M. of the third day. On the fourth day he received 75 micrograms at 10:30 A.M., and on the last day, at 10:40 A.M., he received 100 micrograms of LSD-25.

Subject *C* also participated in two series of experiments. In the first series she received increasing doses of LSD-25 beginning with five micrograms administered at 9:00 A.M. and 5:30 P.M. The following day she received 10 micrograms of the drug at 8:30 A.M. and 5:30 P.M. On the third day she received 20 micrograms of the drug at 9:00 A.M. and 4:30 P.M. On the fourth day, she received 50 micrograms at 10:30 A.M., and on the last day, 75 micrograms at 9:00 A.M. In the second series of experiments the subject received 100 micrograms of BOL-148 at 11:30 A.M. and 6:30 P.M. The next day she received 100 micrograms of BOL-148 at 9:00 A.M. and 10:00 P.M. She received 25 micrograms of LSD-25 at 9:00 P.M. of the third day.

C. PROCEDURE

The subjects had little or no food before receiving the drug which was given orally in 75 cc of tap water. Following ingestion of the drug the subjects responded to a questionnaire $\frac{1}{2}$ hour after receiving the drug and,

in most cases, at four hourly intervals thereafter. The questionnaire contained 47 questions inquiring about the subject's physiological and perceptual state and is reproduced in a previous paper (2). Positive responses ranged from + to +++++, indicating the severity of the symptom present. Subjects also indicated whether normal or not, in the following psychic areas: motor behavior, control, consciousness, concentration, mood, attitude toward environment, orientation, memory, and hallucinations.

D. RESULTS

The number of responses given during each experiment and the specific responses made are summarized here. The results obtained from each subject will be reported separately. Statistical analysis and grouping of the data is not warranted with so few subjects.

1. Subject A

This subject received 100 micrograms of the drug on six successive days and once again five days later. Table 1 indicates the psychic areas in which changes were reported and the total number of times psychic changes were

TABLE 1
NUMBER OF TIMES SUBJECT A REPORTED CERTAIN PSYCHIC CHANGES
(Subject was questioned six times during each of seven experiments with 100 micrograms of LSD-25.)

Area	Number of times changes were reported						
	1*	2	3	4	5	6	11
1. Motor behavior	4	0	0	0	0	0	0
2. Control	0	0	0	0	0	0	0
3. Consciousness	0	0	0	0	0	0	0
4. Concentration	0	0	0	0	0	0	0
5. Mood	0	0	0	0	0	0	0
6. Attitude toward environment	0	0	0	0	0	0	0
7. Orientation	0	0	0	0	0	0	0
8. Memory	0	0	0	0	0	0	0
9. Hallucinations	3	1	0	0	0	0	0
Total	7	1	0	0	0	0	0

*Subject was questioned only five times on this day.

reported during the day. The subject was questioned six times. There were hallucinations on the first two days and changes in motor behavior on the first day. On subsequent days the subject was normal in all areas.

Figure 1 shows the total number of questions receiving positive responses during each question period on each experimental day and the total number of responses made each day.

The subject responded at $\frac{1}{2}$, $1\frac{1}{2}$, $2\frac{1}{2}$, $3\frac{1}{2}$, $4\frac{1}{2}$, and more than $4\frac{1}{2}$ hours after receiving the drug, except on the first day when there was no response during the last interval. The boxed insert on the figure shows that the total number of responses went from 30 to 13, to 15, and to 7 on the

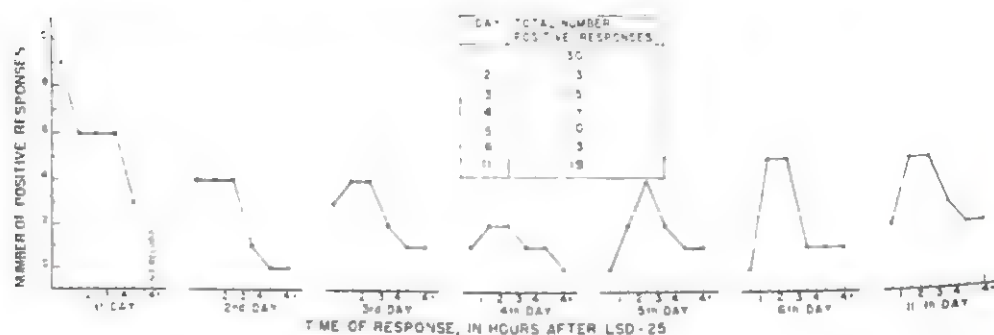


FIGURE 1

Total number of positive questionnaire responses given by Subject A on Section 1 during each of seven experiments. The subject received 100 micrograms of LSD-25 on six successive days and again after five days.

first four days, and then up to 10 on the fifth, and 13 on the sixth. Five days later, when the subject again received 100 micrograms of LSD-25 she gave a total of 19 responses. On the fifth day a decreased response occurred only during the first $\frac{1}{2}$ hour. The maximum number of responses given during the last three intervals was two on all but the first day. On the eleventh day the number of responses given was greater than on the second day (except for the first $\frac{1}{2}$ hour) but not as great as on the first day.

The hourly responses to each question on each experimental day are graphed in Figure 2. Where the same response was given on more than one day, the curves appear beneath each other. Only those questions receiving at least one positive response appear on the graph. Two questions received positive responses during every experiment, but with varying frequency: (No. 35) "Is your eyesight blurred?" and (No. 42) "Do you tremble inside?" Blurred eyesight was least frequent on the sixth day, while inner trembling was least prominent on the fourth day. The greatest response was given on the first and seventh days. Some symptoms, awareness of heartbeat (No. 19), heartbeat faster than usual (No. 20), and moist palms (No. 24), were reported only during the first experiment. Others, lips drawn back as if smiling (No. 12), diplopia (No. 37), and shapes and colors altered (No. 38), reported during the first few days disappeared completely in the middle days and reappeared on the sixth day. One symptom, objective vertigo (No. 14), was reported only on the first and fifth days.

Two symptoms, headache (No. 13) and dizziness (No. 15), appeared only in the later experiments. The magnitude of the responses seemed to vary with their frequency.

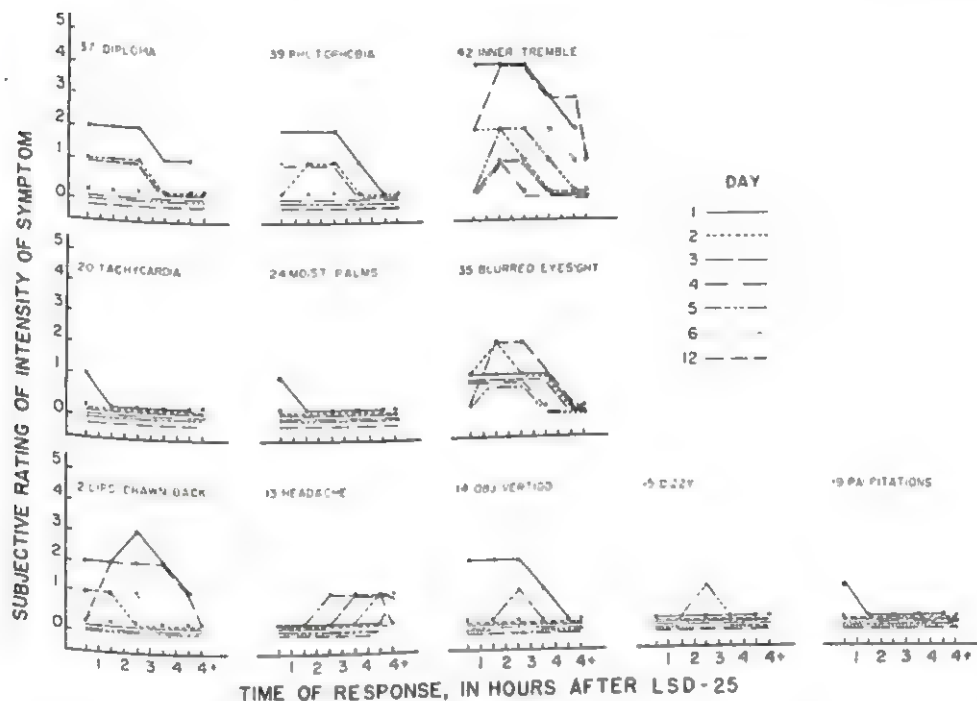


FIGURE 2

Subjective rating of intensity of each symptom reported by Subject A on Section I of the questionnaire. The subject received 100 micrograms of LSD-25 on six successive days and again after five days.

2. Subject B

This subject received 100 micrograms of LSD-25 on three successive days. Table 2 shows that during the first two there were changes in all psychic areas but one: no hallucinations occurred. On the third day there were alterations only in consciousness, concentration, and mood.

Figure 3 gives the number of positive responses made each day. There was a total of 64, 37, and 21 positive responses on the three days, respectively. Comparison of the time curves for these days clearly demonstrates the daily decline. Specific hourly responses made to each question appear in Figure 4. Only those questions receiving at least one positive response are graphed. The symptoms reported on all three days were: feeling of choking (No. 3), headache (No. 13), moist palms (No. 24), blurred eyesight

(No. 35), difficulty in focusing vision (No. 36), inner trembling (No. 42), weakness (No. 43), and fatigue (No. 44). Some symptoms were reported only on the first two days: decreased salivation (No. 5), objective vertigo (No. 14), dizziness (No. 15), unsteadiness (No. 16), sweating (No. 21),

TABLE 2

NUMBER OF TIMES SUBJECT B REPORTED CERTAIN PSYCHIC CHANGES

(Subject was questioned five times during each of three experiments with 100 micrograms of LSD-25.)

Area	Number of times changes were reported		
	1	2	3
1. Motor behavior	3	2	0
2. Control	3	3	0
3. Consciousness	2	3	1
4. Concentration	3	3	2
5. Mood	4	2	2
6. Attitude toward environment	3	2	0
7. Orientation	3	3	0
8. Memory	2	3	0
9. Hallucinations	0	0	0
Total	23	21	5

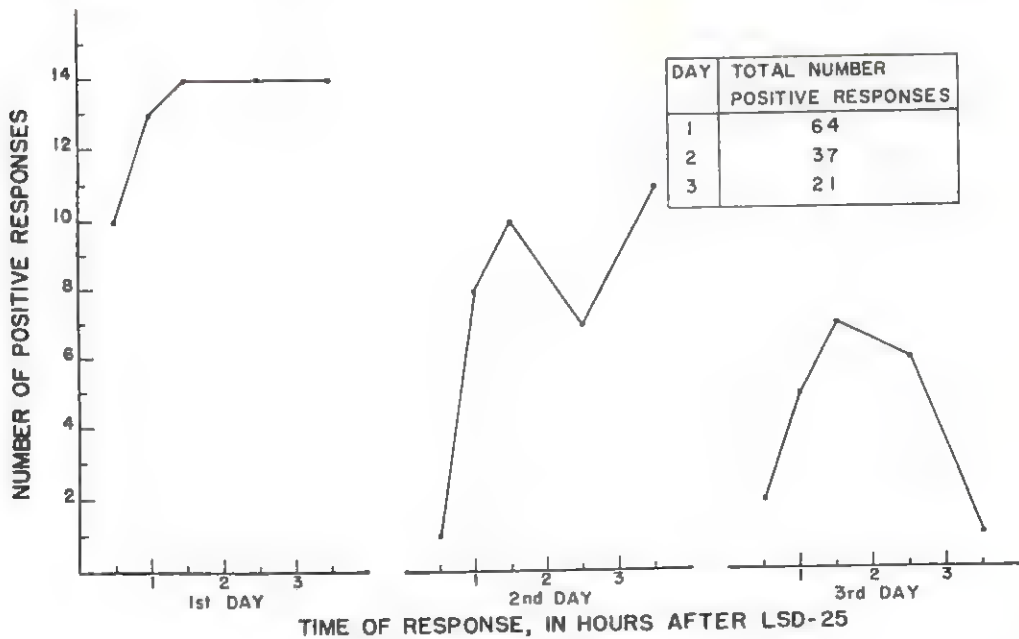


FIGURE 3

Total number of positive questionnaire responses given by Subject B on Section I during each of three experiments. The subject received 100 micrograms of LSD-25 on three successive days.

warmth (No. 22), photophobia (No. 39), apparent closeness of objects (No. 40), and a dream-like feeling (No. 46). On the first day, only, the subject reported dry taste in mouth (No. 8), lips drawn back as if smiling (No. 12), awareness of heartbeat (No. 19), pressure in ears (No. 32),

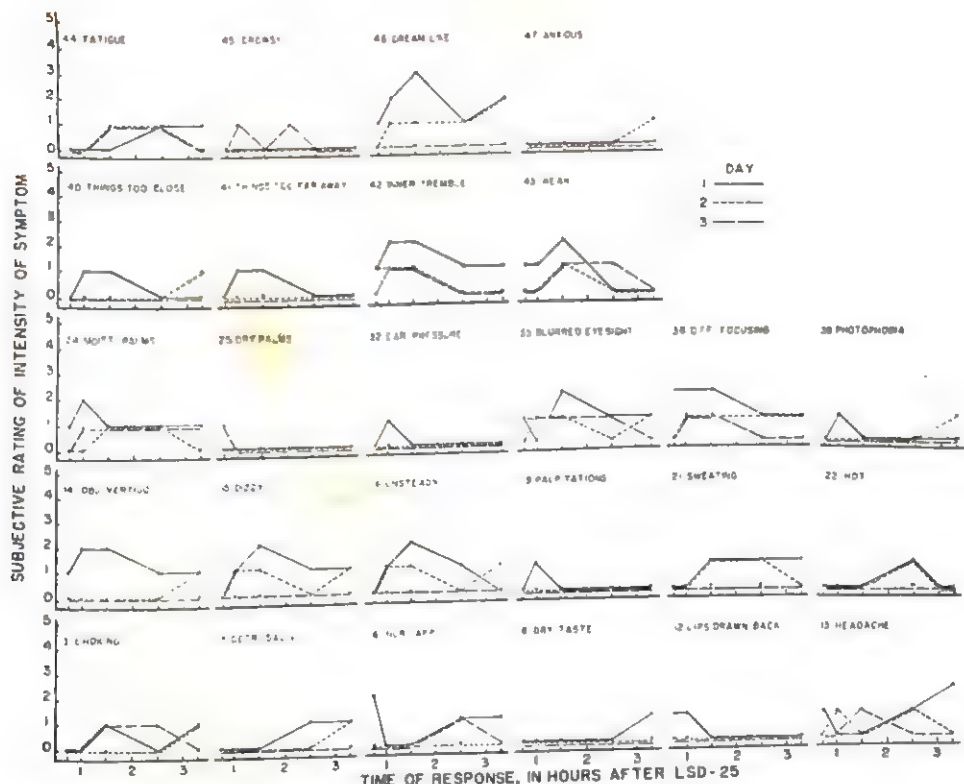


FIGURE 4

Subjective rating of intensity of each symptom reported by Subject B on Section I of the questionnaire. The subject received 100 micrograms of LSD-25 on three successive days.

and things seemed too far away (No. 41). Irregularly appearing responses were, increased appetite (No. 6), dry palms (No. 25), drowsiness (No. 45), and anxiety (No. 47). The responses generally appeared with less frequency and intensity on successive days.

In the second series of experiments for this subject he received increasing doses of LSD-25 on five successive days. The psychic changes which occurred are enumerated in Table 3. Not until the fourth day, when 75 micrograms of the drug was administered, was there a psychic change; the subject re-

ported a slight decrease in concentration. After 100 micrograms of LSD-25 a decrease in consciousness, slight confusion, and a feeling of remoteness from the environment were reported.

TABLE 3
NUMBER OF TIMES SUBJECT *B* REPORTED CERTAIN PSYCHIC CHANGES
(Subject was questioned six times during each of seven experiments with increasing doses of LSD-25.)

Area	Number of times changes were reported Day and dose						
	1-5 μ g.	1a-5 μ g.	2-10 μ g.	3-25 μ g.	3a-25 μ g.	4-75 μ g.	5-100 μ g.
1. Motor behavior	0	0	0	0	0	0	0
2. Control	0	0	0	0	0	0	0
3. Consciousness	0	0	0	0	0	0	1
4. Concentration	0	0	0	0	0	2	1
5. Mood	0	0	0	0	0	0	0
6. Attitude toward environment	0	0	0	0	0	0	1
7. Orientation	0	0	0	0	0	0	0
8. Memory	0	0	0	0	0	0	0
9. Hallucinations	0	0	0	0	0	0	0
Total	0	0	0	0	0	2	3

Figure 5 shows the number of positive responses during each experiment. During four of the seven experiments there were none. After 75 and 100 micrograms the subject gave six positive responses. He did indicate that 2½ hours after receiving 10 micrograms of the drug he could detect its presence. The particular responses made are given in Figure 6. Under 25 micrograms of the drug, given in the morning, he reported drowsiness (No. 45) and awareness of the presence of LSD-25. Under 75 micrograms the subject reported difficulty in focusing his vision (No. 36), photophobia (No.

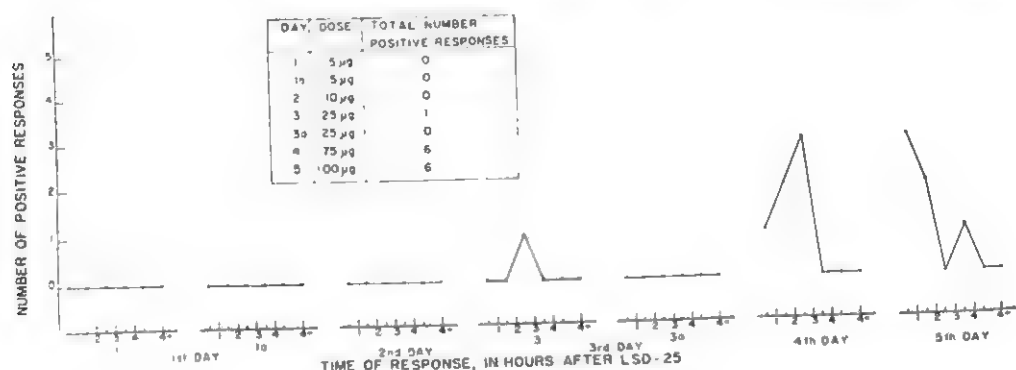


FIGURE 5

Total number of positive questionnaire responses given by Subject *B* on Section I during each of seven experiments. The subject received increasing doses of LSD-25 on successive days.

39), inner trembling (No. 42), and drowsiness (No. 45). After 100 micrograms he reported a feeling of choking (No. 3), unsteadiness (No. 16), blurred eyesight (No. 35), inner trembling (No. 42), fatigue (No. 44), and drowsiness (No. 45).

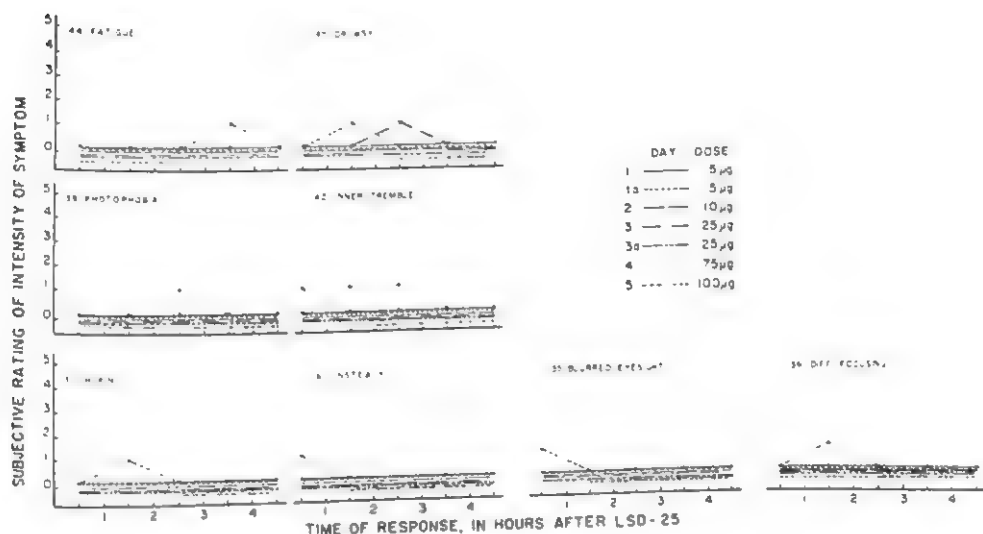


FIGURE 6

Subjective rating of intensity of each symptom reported by Subject B on Section I of the questionnaire. The subject received increasing doses of LSD-25 on successive days.

3. Subject C

Subject C also received increasing doses of LSD-25. The changes which occurred in the psychic areas appear in Table 4. No changes were reported until the fourth day, when the subject received 50 micrograms of the drug and reported impairment in motor behavior, control, concentration, and a detached attitude toward the environment. Under 100 micrograms she reported impaired coordination, decreased ability to concentrate, and a detached attitude toward the environment.

Figure 7 shows the total number of positive responses at each hour and each day. Under 5 and 10 micrograms of LSD-25 the subject gave no positive responses at either administration. Under 20 micrograms of the drug, taken in the morning, she gave a total of six responses. The same dose given in the evening evoked no positive response. Under 50 micrograms of LSD-25 on the following day, there was a total of 11 responses and under 75 micrograms there were 14.

The specific responses made under each dose are graphed in Figure 8. After the first dose of 20 micrograms of LSD-25 Subject *C* reported slight nausea (No. 2), dizziness (No. 15), heaviness of hands and feet (No. 30),

TABLE 4
NUMBER OF TIMES SUBJECT *C* REPORTED CERTAIN PSYCHIC CHANGES
(Subject was questioned five times during each of eight experiments with increasing doses of LSD-25.)

Area	Number of times changes were reported							
	Day and Dose							
	1-5 μ g.	1a-5 μ g.	2-10 μ g.	2a-10 μ g.	3-20 μ g.	3a-20 μ g.	4-50 μ g.	5-75 μ g.
1. Motor behavior	0	0	0	0	0	0	1	1
2. Control	0	0	0	0	0	0	1	0
3. Consciousness	0	0	0	0	0	0	0	0
4. Concentration	0	0	0	0	0	0	1	3
5. Mood	0	0	0	0	0	0	0	0
6. Attitude toward environment	0	0	0	0	0	0	1	2
7. Orientation	0	0	0	0	0	0	0	0
8. Memory	0	0	0	0	0	0	0	0
9. Hallucinations	0	0	0	0	0	0	0	0
Total	0	0	0	0	0	0	4	6

and weakness (No. 43). Under the second dose of 20 micrograms there were no positive responses. The symptoms reported after 50 micrograms of the drug were: unsteadiness (No. 16), peculiar feeling in hands and feet (No. 29), heaviness of hands and feet (No. 30), apparent alterations

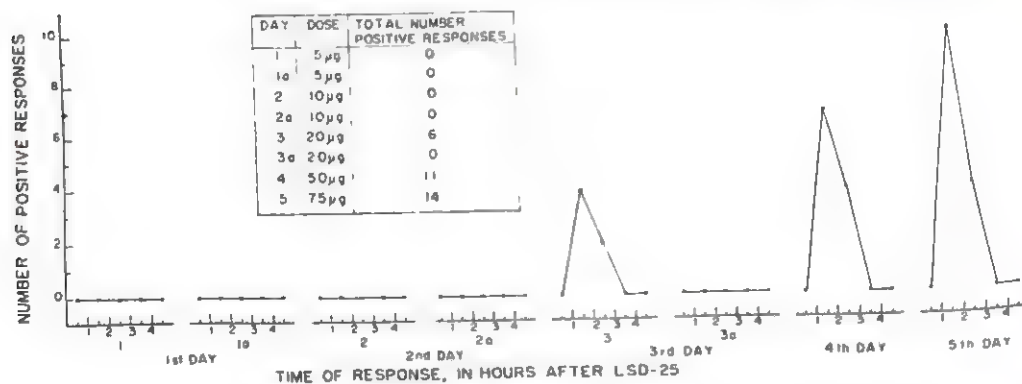


FIGURE 7

Total number of positive questionnaire responses given by Subject *C* on Section I during each of eight experiments. The subject received increasing doses of LSD-25 on successive days.

in shapes and colors (No. 38), inner trembling (No. 42), weakness (No. 43), and a dream-like feeling (No. 46).

Under the highest dose (75 micrograms) some symptoms appearing under 50 micrograms were reported: unsteadiness (No. 16), peculiar feeling

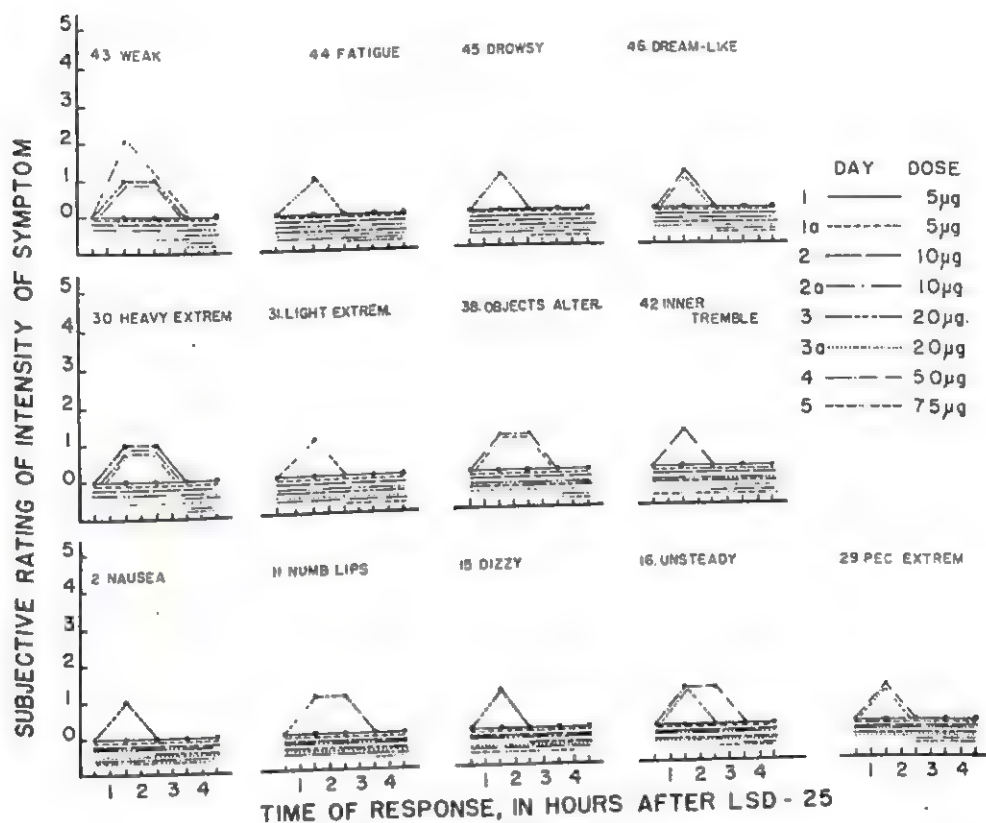


FIGURE 8

Subjective rating of intensity of each symptom reported by Subject C on Section I of the questionnaire. The subject received increasing doses of LSD-25 on successive days.

in hands and feet (No. 29), heaviness of hands and feet (No. 30) apparent alterations in shapes and colors (No. 38), weakness (No. 44), and dream-like feelings (No. 46). In addition to these responses, numb lips (No. 11), lightness of hands and feet (No. 31), fatigue (No. 44), and drowsiness (No. 45) were indicated. Most responses were made 1½ hours and some were made 2½ hours after the drug.

In the second series of experiments this subject received 100 micrograms of BOL-148 four times; then 25 micrograms of LSD-25. During the first

experiment there were changes in motor behavior, control, and concentration (see Table 5). There were no psychic changes during the subsequent experiments.

TABLE 5

NUMBER OF TIMES SUBJECT C REPORTED CERTAIN PSYCHIC CHANGES

(Subject was questioned six times during each of four experiments with 100 micrograms of BOL-148 and a fifth experiment with 25 micrograms of LSD-25.)

Area	Number of times changes were reported				
	Day				
	1	1a	2	2a*	3
1. Motor behavior	3	0	0	0	0
2. Control	2	0	0	0	0
3. Consciousness	0	0	0	0	0
4. Concentration	3	0	0	0	0
5. Mood	0	0	0	0	0
6. Attitude toward environment	0	0	0	0	0
7. Orientation	0	0	0	0	0
8. Memory	0	0	0	0	0
9. Hallucinations	0	0	0	0	0
Total	8	0	0	0	0

*Subject was questioned only three times on this day.

Figure 9 shows that there were five positive responses following the first administration of BOL-148, nine after the second, five after the third, and none after the fourth. Under LSD-25 there was a total of 12 positive responses. No positive responses were given during the first $\frac{1}{2}$ hour.

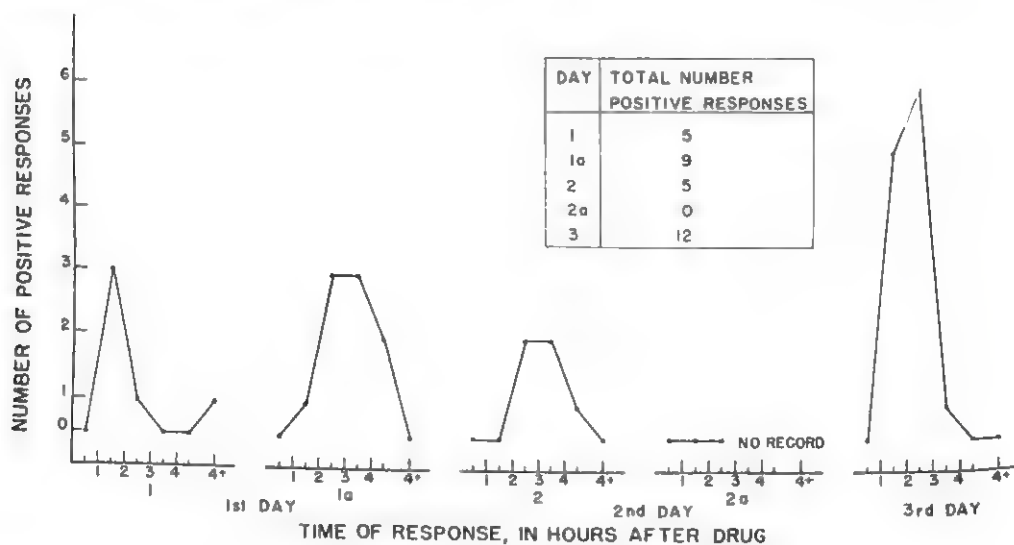


FIGURE 9

Total number of positive questionnaire responses given by Subject C on Section I during each of five experiments. The subject received 100 micrograms of BOL-148 twice daily for two days and 25 micrograms of LSD-25 on the third day.

Figure 10 indicates which questions received the positive responses during each experiment. After the first experiment with BOL-148 the subject reported a headache (No. 13), dizziness (No. 15), funny feelings on the skin (No. 28), and inner trembling (No. 42). Under the second dose

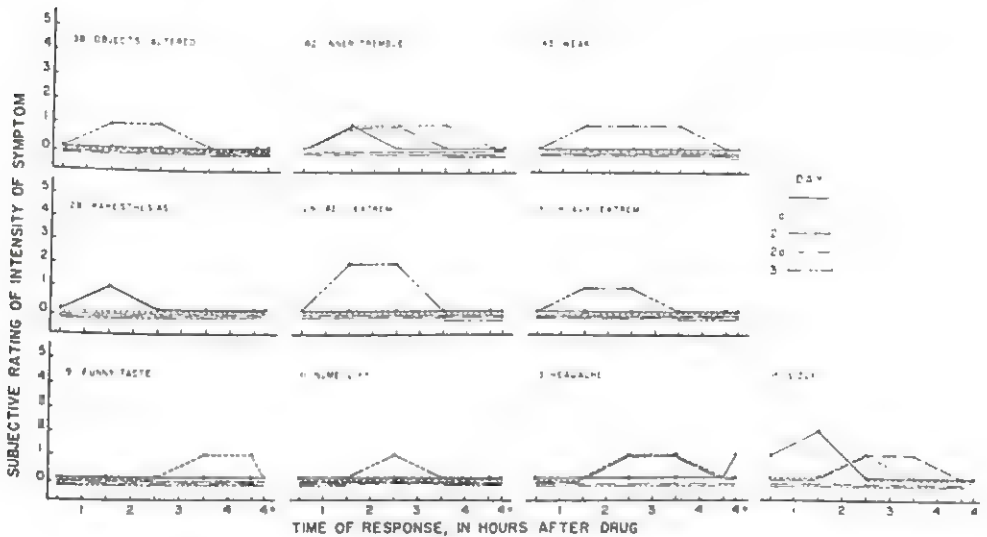


FIGURE 10

Subjective rating of intensity of each symptom reported by Subject C on Section I of the questionnaire. The subject received 100 micrograms of BOL-148 twice daily for two days and 25 micrograms of LSD-25 on the third day.

of BOL-148 in the evening of the first day she again reported a headache (No. 13), dizziness (No. 15), and inner trembling (No. 42), and also a funny taste in her mouth (No. 9). After BOL-148 given in the morning of the following day, headache (No. 13) and dizziness (No. 15) were reported. That evening she gave no positive responses to 100 micrograms of BOL-148. The next day the subject received 25 micrograms of LSD-25 and reported a total of six different symptoms: numb lips (No. 11), peculiar feeling in hands and feet (No. 29), heaviness in hands and feet (No. 30), apparent alterations in shapes and colors (No. 38), inner trembling (No. 42), and weakness (No. 43). All positive responses were given $2\frac{1}{2}$ hours after receiving the drug and all symptoms but numb lips (No. 11) were also reported at $1\frac{1}{2}$ hours.

Further study of the figures shows the hourly changes in severity of response to each question.

E. DISCUSSION

The pattern of rapid development of tolerance, and its correspondingly rapid decline as reported by Isbell (6) is not typical of either immuno-type mechanisms or detoxification mechanisms. To provide possible insight into the autodote phenomenon, a mechanism has been formulated which involves the concept of a tolerance factor stoichiometrically related to the drug itself, and formed as a direct consequence of its physiological action.

Isbell (6) reports that tolerance is essentially completely lost in three days. Our data on rate of loss of tolerance are meager, and confined to Subject *A*. In this subject considerable tolerance remained five days after the last administration of the drug. This is shown by the experiment wherein tolerance had been established by the administration of six 100-microgram doses spaced 24 hours apart, and tested by a seventh dose administered five days later (on the eleventh day). Indirect evidence from the analysis of the data on Subjects *B* and *C* indicates that their patterns may be like that of Isbell's subjects, even though they are psychologically different in some other respects.

In any event, studies of a wide variety of individuals to determine whether characteristic patterns with respect to rate of development and loss of tolerance can be established and reproduced are indicated. The mechanism to be proposed offers a way to correlate the data in terms of a single rate constant (k_4) for any one series of experiments.

While this is undoubtedly an over-simplification of a very complex mechanism, the rate constant appears to define a key factor in the systems.

The extremely low dosage at which initial reaction is obtained is of the order of that required for systemic reactions of histamine and epinephrine. It points to an intrusion of LSD-25 into psychic reactions as an analog of physiologic substances normally involved in these reactions; the unique feature of the action is that tolerance is so quickly established, and yet so rapidly lost. In this respect it differs in kind from histamine and epinephrine where tolerance to repeated administration has not been successfully demonstrated except over much longer periods, if at all. Yet caution is necessary in formulating any theory, for while LSD-25 may be a physiologic substance involved in the neurological phenomena or related to such substances, it is possible that the method of administration sets up an artificial situation. By distributing the material through the ordinary channels it comes into contact with tissues where normally it might have no metabolic function, but still could have metabolic consequences. This would set in motion a detoxification mechanism wherein the LSD-25 is destroyed or neutralized (eliminated) before it can reach the centers of its primary reaction. In other

words, the establishment of tolerance could be an artifact of the method of administration not intimately connected with the mechanism of the psychic reactions induced by the drug.

In the speculations which follow, we are taking the view that the establishment of tolerance and the rapid loss of tolerance are part of a unified mechanism which also involves the psychic actions of the drug. The motivating factor in this choice of routes is that the pattern of rapid establishment and loss of tolerance is unique and not typical of detoxification mechanisms.

Finally, the question arises whether all the symptoms observed are due to a primary reaction of LSD-25 at one particular neurological site, or to several independent reactions at different sites.

The fact that Isbell (6) obtains complete tolerance (zero reaction by the questionnaire) after several weeks of administration can be used to support the notion of one primary reaction being involved. We have not obtained complete tolerance to larger doses of LSD-25, but our periods of administration are shorter, and zero reaction might have been obtained if administration had been continued over longer periods for Subjects *B* and *C*.

For Subject *A* it appears that complete tolerance to large doses (100 micrograms) cannot be obtained, even though partial tolerance is very quickly established and relatively slowly lost. The mechanism to be presented is sufficiently flexible to cover cases of complete and incomplete tolerance without invoking more than one primary reaction step.

1. *Stoichiometric Relations for Tolerance*

In the case of Subject *A*, there is almost a 1/1 correspondence between LSD-25 administered and the development of tolerance. Thus, 100 micrograms given on Day 1 protects the subject almost completely against a 100-microgram dose on Day 2, even though a reaction greater than a threshold reaction is observed.

In the case of Subject *B*, there is approximately a $\frac{1}{2}$ correspondence between LSD-25 administered and the development of tolerance. Thus, 100 micrograms given on Day 1 protects the subject against about $\frac{1}{2}$ of a 100-microgram dose on Day 2. Also, a total of 45 micrograms given in four doses on Days 1, 2, and 3 protects completely against 25 micrograms given 12 hours later on Day 3.

Similar correspondence exists in Subject *C*, and in Isbell's subjects. Obviously, since tolerance is being rapidly lost as well as established, only in cases where the rate of loss is relatively slow can the ideal stoichiometric 1/1 correspondence be approached. In all actual cases it would have to be less

than 1/1. Subject *A* does show a slow loss of tolerance, and the nearly 1:1 correspondence between initial dosage and tolerance 24 hours afterward ties in with the slow loss, and offers confirmation for the theory.

2. Mechanism

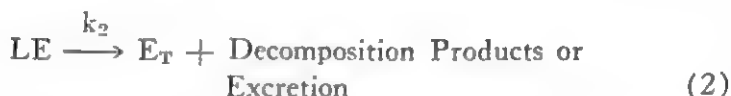
The mechanism to be suggested for the phenomenon of tolerance has four essential steps. In Step 1, LSD-25 (designated by *L*) interacts with a neuro-metabolic system, *E*, to cause symptoms, and becomes modified by combination with a component of this system and is converted to *LE*. Reaction 1 (see below) is reversible so that as long as *LE* is present, some reaction of LSD-25 might be observed.

LE is labile and in Step 2, *LE* splits into *E_T* and other products which are eliminated. *E_T* is the key substance in the mechanism, and designates the tolerance factor.

When *E_T* comes in contact with LSD-25, it reacts with it to form *LE*. This is Step 3 in the mechanism, and it follows that when tolerance is established LSD-25 is preferentially reacting with *E_T* compared with the metabolic reaction which causes symptoms. It is postulated that *LE* is reformed because this allows *E_T* to be built up in time by repeated administration of the drug.

Since tolerance disappears in time, a fourth step is postulated wherein *E_T* is eliminated.

These four steps are represented by the equations below. The *k*'s over the arrows are the rate constants for the individual steps.



All four reactions proceed while symptoms are occurring, but when the effect of the drug wears off only Reaction 4 continues to go on. The psychic reactions appear to be over in hours while the elimination of tolerance takes several days. It follows, therefore, as a first approximation, that when the

psychic reaction is over, essentially all of the LSD-25 which entered the site of reaction has been converted to E_T . Also, it is obvious that the intensity of the reaction to the next administration of LSD-25 is primarily determined by the rate of loss of E_T by Reaction 4. Initially, enough E_T is present to protect against an approximately equal dose of LSD-25, but the longer the period between administrations (for a given value of k_4), the greater will be the reaction to the next dose.

Postulating that Reaction 1 is reversible brings in the feature of residual response to LSD-25 after tolerance is established. A steady state between Reactions 1 and 5 reduces the effective concentration of L to a low level. If this is below the threshold for the individual concerned zero reaction will be obtained. On the other hand, the steady state concentration in other individuals might be above their threshold level and some symptoms occur. This seems to correspond to the situation with Subject *A*.

3. Quantitative Considerations

The quantitative considerations are based on the idea that Reaction 4 is of the first order with respect to E_T , and therefore that in any given time period between administrations a constant percentage of E_T will be lost irrespective of the initial amount present. Further, the loss of E_T with time will follow the equation:

$$\log \frac{A}{A-x} = k_4 t \quad (6)$$

where A is the initial amount of E_T present and x is the amount of E_T lost in time, t . To illustrate the way x varies with k_4 Table 6 was computed from Equation (7) for a value of t of 24 hours.

TABLE 6
VALUES OF k_4 AS A FUNCTION OF LOSS OF TOLERANCE

Per cent loss of tolerance in 24 hours, x	k_4 , hours ⁻¹	$1/k_4$, hours
10	0.0043	232
20	0.0095	105
40	0.0210	47.6
60	0.0378	26.5
80	0.0662	15.1
90	0.0905	11.0
95	0.219	5.57
99	0.288	3.47

The units of k_4 are reciprocal time. The reciprocal of k_4 (Column 3) represents the time in hours at which

$$\ln \frac{100 - x}{100} = 1 \quad (7)$$

or at which $x = 63.5$ per cent loss of tolerance.

Computations were based on a repetitive dose of LSD-25 of 100 micrograms every 24 hours. Three values of k_4 were used—one corresponding to 50 per cent loss of E_T in 24 hours, one to 40 per cent loss, and one to 60 per cent loss. Results are given in Table 7.

TABLE 7
 E_T REMAINING ON SUCCESSIVE ADMINISTRATIONS OF LSD-25, WITH EACH OF THREE VALUES OF k_4

Admin. of LSD-25	1st	2nd	3rd	4th	5th	6th	
Case I— k_4 Corresponds to 50 per cent							
Loss of E_T in 24 hours							
E_T Remaining*	0	50	75	87.5	93.75	97.88	
Effective Dose of LSD-25**	100	50	25	12.5	6.25	2.12	
Case II— k_4 Corresponds to 40 per cent							
Loss of E_T in 24 hours							
E_T Remaining*	0	60	96	117.6	130.6	138.4	
Effective Dose of LSD-25**	100	40	4	— 17.6	— 30.6	— 38.4	
Case III— k_4 Corresponds to 60 per cent							
Loss of E_T in 24 hours							
Admin. of LSD-25	1st	2nd	3rd	4th	5th	6th	7th
E_T Remaining*	0	40	56	62.4	65	66	66.4
Effective Dose of LSD-25**	100	60	44	37.6	35	34	33.6

*Twenty-four hours after most recent administration of LSD-25

**For this administration of LSD-25.

The three cases tabulated illustrate the critical effect of the value of k_4 on the build up of tolerance to the drug. A value of k_4 corresponding to 50 per cent loss of E_T allows build up of complete tolerance. If k_4 increases to 60 per cent, then only incomplete protection (tolerance) is obtained, and a 100-microgram dose will, after seven successive daily administrations, act like about 1/3 of this dose given initially. On the other hand, values of k_4 corresponding to less than 50 per cent loss allow build up of tolerance to more than the initial dose. Thus, for k_4 corresponding to 40 per cent loss per 24 hours a tolerance to 138.4 micrograms of the drug will build up in six days.

Another factor of interest is the effect of time between administrations.

An individual corresponding to Case III (60 per cent loss of E_T in 24 hours) would act like Case I (50 per cent loss) by reducing the time between administrations to 18 hours. This is computed from equation (6) as follows:

$$\begin{aligned} A &= 100 \\ x &= 60 \\ t &= 24 \\ \therefore \log \frac{100}{100-60} &= 24k_4 \\ \text{or } k_4 &= \frac{\log \left(\frac{100}{40} \right)}{24} = \frac{\log 2.5}{24} \\ &= \frac{0.398}{24} = 0.0167 \text{ hours}^{-1} \end{aligned}$$

To calculate time for 50 per cent loss of E_T insert the value of k_4 , 0.0167, into the equation and solve for t :

$$\begin{aligned} t &= \frac{1}{0.0167} \log \frac{100}{100-50} \\ &= \frac{\log 2}{0.0167} = \frac{0.300}{0.0167} = 18.0 \text{ hours.} \end{aligned}$$

4. Application of Model to Specific Cases

a. *Subject A.* A value of k_4 corresponding to a 25 per cent loss of $E_T/24$ hours gives a pattern which is in reasonable agreement with the data. It predicts that the reactions on Days 2 and 11 should be roughly equivalent and slightly greater than those on Days 3, 4, 5, and 6.

TABLE 8
 E_T REMAINING ON SEVEN SUCCESSIVE ADMINISTRATIONS OF LSD-25 AND EFFECTIVE DOSE PREDICTED
(k_4 corresponds to 25 per cent loss/24 hours)

Day	1	2	3	4	5	6	7	8	9	10	11
Dose	100	100	100	100	100	100	—	—	—	—	100
E_T^*	—	75	131	173	204	228	246	185	139	104	78
Effective Dose	100	25	— 31	— 73	—104	—128	—	—	—	—	22

*Remaining from previous administration.

Note that after Day 2 complete tolerance to 100 micrograms is predicted, building up to a tolerance to almost 250 micrograms at the beginning of Day 6. It would be interesting to test this experimentally. Also, reactions on Days 3, 4, 5, and 6 should be equivalent.

b. Subject B. The two series on Subject *B* cannot be reconciled with the theory. In Series 1 wherein the 100 micrograms of the drug was administered in three equal doses 24 hours apart, the data correspond to a k value of 50 per cent, which predicts (Table 9).

TABLE 9
PREDICTED EFFECTIVE DOSE OF LSD-25
(k_1 corresponds to 50 per cent loss of E_T 24 hours)

Day	Effective Dose
1	100
2	50
3	25

In Series 2, a smaller rate constant of the order of 30 per cent loss 24 hours would have to be used to explain the data. This anomaly need not be explained with the present state of our knowledge, since sufficient data are not available to establish that either series are reproducible for a given individual. If this were established—then the model would have to be abandoned or modified.

In any event, in Series 2, this subject was made almost completely tolerant to a 75-microgram dose by the previous administration of only 70 micrograms in five doses over a three-day period.

c. Subject C. Computations based on a rate constant corresponding to 50 per cent loss per 24 hours gives a pattern similar to that observed in Subject *C*. For a 12-hour period the loss would be 29 per cent according to Equation (6). Results of computation with this value of k_1 are given in Table 10.

TABLE 10
 E_T REMAINING WITH SUCCESSIVELY INCREASING DOSES OF LSD-25 AND EFFECTIVE DOSE
PREDICTED
(k_1 corresponds to 50 per cent loss/24 hours)

Subject <i>C</i>	1st	2nd	3rd	4th	5th	6th	7th	8th
Dose, micrograms	5	5	10	10	20.0	20.0	50	75
Time elapsed since last administration, hours	—	12	12	12	12	12	12	24
E_T remaining from last administration	0	3.5	6.0	11.3	15.1	24.8	31.3	40.7
Effective dose, computed	5	1.5	4.0	— 1.3	4.9	— 4.8	18.7	34.3
Effective dose, found	*	*	*	*	*	*	**	***

*No reaction.

**Positive reaction somewhat less than that to 25 micrograms given in isolated dose.

***Positive reaction approximately like that found for 25 micrograms in isolated dose.

d. *Isbell's Data.* Isbell's (6) subjects, taken as a group, show correspondence with a k value of approximately 50 per cent loss per 24 hours (29 per cent loss in 12 hours). Carrying the computations through for his series, the following is obtained (Table 11), which predicts a 25-microgram response to the 75-microgram dose given in the seventh administration. By continuing to administer 75 micrograms, 24 hours a build up of tolerance to a plateau value of 75 micrograms is predicted for this value of k_4 . Having established this, if administration of LSD-25 is stopped, tolerance would decline according to a k_4 value of 50 per cent/24 hours (Table 12).

TABLE 11
E_T REMAINING WITH SUCCESSIVELY INCREASING DOSES OF LSD-25 AND EFFECTIVE DOSE
PREDICTED
(k_4 corresponds to 50 per cent 24 hours)

Isbell's subjects	1st	2nd	3rd	4th	5th	6th	7th
Dose, micrograms	10	10	20	20	30	30	75
Time elapsed since previous administration, hours	—	12	12	12	12	12	12
E _T remaining from previous administration	0	7.1	12.1	22.8	31.8	41.9	50.1
Effective dose	10	2.9	7.9	— 2.8	— 1.8	11.9	24.9

TABLE 12
PREDICTED LOSS OF TOLERANCE WHEN LSD-25 IS NO LONGER GIVEN

Day	Tolerance	Effective dose
0	75	0
1	37.5	37.5
2	18.7	56.5
3	9.3	65.3

A value of k_4 of 50 per cent thus predicts that tolerance would be lost in approximately three days.

5. *Application of Mechanism to Schizophrenia*

In offering a mechanism for the development and loss of tolerance to LSD-25, the question naturally arises whether this mechanism contains some clue to the nature of schizophrenia. Such a question contains at least three assumptions; namely: (a) That there is a substance analogous to LSD-25 involved in schizophrenia, and designated here as P . (b) That P has some normal function in the physiology of the emotional processes, but the metabolism of P is disturbed in schizophrenia. (c) That the substance P is regulated by a mechanism like the one proposed here for the development and

loss of tolerance to LSD-25, and that a breakdown of this mechanism is involved in P getting out of control.

Thus, we would have the mechanism:



where the normal function of P might be involved, for instance, in the accommodation of the organism to stress situations. E_T would be the ballast against emotional explosion while the reaction k_5 would be the "safety valve" to bring about the necessary emotional response. Without such a safety valve E_T could act to eliminate P before the organism could make the necessary response.

There are data on the action of LSD-25 on schizophrenics. For instance, Cholden (3) found that schizophrenics become tolerant to LSD-25 and lose that tolerance in a manner approximately parallel to that of normals. Thus, the k_4 value (rate of elimination of E_T) is not outstandingly different in schizophrenics, compared with normals.

Further studies on normals and schizophrenics might result in the demonstration of differences in k_4 . This would be important for classification and diagnosis. It would not seem to be the key to the problem of schizophrenia.

A more tenable hypothesis is that while schizophrenics show a relatively normal tolerance pattern, this pattern is ineffective in protecting against psychic reactions because of an abnormally high value of k_5 . This would lead to a steady state concentration of the P factor above the value required to induce psychotic reactions even though large amounts of the tolerance factor, E_T are present. Building up E_T to larger values would prevent the reactions to the massive doses of LSD-25 as observed by Cholden (3), but the effect of small doses would not be observable in the psychotic patient over their usual symptoms.

Another rate constant in the mechanism which would have a profound effect on the nature of the response to P is k_2 , the conversion of PE to E_T . If PE is reversing into P and E , and thus causing recurring symptoms, the

chain is broken only by Reaction 2, by which means PE goes to E_T , the tolerance factor. A subject with a high k_3 , and a high k_2 would show bursts of psychotic behavior of short duration, within the approximately normal range. A low value of k_2 would prolong the psychotic state induced by the "secretion" of P .

It would be of interest to test this hypothesis by experiments in which the duration of the psychic reactions induced by LSD-25 are compared in normals and schizophrenics.

6. *Application of Theory to Biochemical Studies*

Experimental work is being instituted (1) to detect agents in the urines of clinically schizophrenic patients which might be the cause of clinical schizophrenia in man (P substance). The ideas developed above indicate that anti- P substances (E_T) should also be sought in such urines and in the urines of non-schizophrenics.

Referring back to the mechanism, if P is formed ("secreted") only at the site of its primary action it would not be eliminated as P but would end up as decomposition products of E_T , the tolerance factor, formed by Reaction 4. E_T might also be directly eliminated from the sites of neurological action and be found in the urine.

Another possible chemical clue in the urines would be the decomposition products of PE formed in Reaction 2, wherein PE is converted to E_T .

F. SUMMARY AND CONCLUSIONS

1. Three subjects with considerable experience in LSD-25 experiments were tested with LSD-25 on six and three consecutive days, respectively. The response to a questionnaire diminished rapidly from day to day. On the fifth and sixth days an increase appeared. Five days after the sixth day the response to 100 micrograms of LSD-25 by Subject A was not as great as on the first day but was greater than on the other days. The lowest number of responses resembled a 25-microgram response. The symptoms reported at that time were more typical of high dose symptoms but appeared less frequently.

2. Subjects B and C received increasing doses of LSD-25 on successive days, beginning with doses of five micrograms and gradually increasing to 100 and 75 micrograms, respectively. Subjects exhibited marked tolerance. Their responses to the highest doses were less than their usual responses to a 25-microgram dose of LSD-25.

3. Subject C demonstrated tolerance to 100 micrograms of BOL-148

given twice daily for two successive days, but showed only slight, if any, tolerance to 25 micrograms of LSD-25 given the following day.

4. All of the foregoing data were coordinated with the theory of the development and loss of tolerance to LSD-25. The mechanism suggested has four essential steps:

a. LSD-25 interacts with a neuro-metabolic system, E , to cause symptoms, and becomes modified, and is converted to LE . This is a reversible reaction and as long as LE is present some LSD-25 reaction might be observed.

b. LE splits into E_T (the tolerance factor) and other products which are eliminated.

c. When E_T comes into contact with LSD-25 it forms LE .

d. With the loss of tolerance in time, E_T is eliminated as a decomposition product or excretion. The loss of E_T will follow the equation:

$$\log \frac{A}{A-x} = k_4 t$$

where A is the initial amount of E_T present and x is the amount of E_T lost in time, t .

5. The theory developed was discussed in relation to the data obtained. Application of the formula to the data demonstrated the relationship between the predicted effective dose of LSD-25 and the effective dose found. Further experimentation was suggested to verify certain predictions.

6. The theory for development and loss of tolerance to LSD-25 was adapted to a theory pertaining to the nature of schizophrenia. A substance P , analogous to LSD-25, was suggested as giving rise to the mechanisms of tolerance which are either lost or altered during clinically psychotic reactions.

7. It was proposed that both the P substance and anti- P substances (E_T) should be sought in the urines of clinically schizophrenic patients, and that anti- P substances might also be found in non-schizophrenic urines.

REFERENCES

1. ABRAMSON, H. A., & EVANS, L. T. Lysergic acid diethylamide (LSD-25): II. Psychobiological effects on the Siamese fighting fish. *Science*, 1954, **120**, 990-991.
2. ABRAMSON, H. A., JARVIK, M. E., KAUFMAN, M. R., KORNETSKY, C., LEVINE, A., & WAGNER, M. Lysergic acid diethylamide (LSD-25): I. Physiological and perceptual responses. *J. of Psych.*, 1955, **39**, 3-60.
3. CHOLDEN, L., KURLAND, A., & SAVAGE, C. Clinical reactions and tolerance to LSD-25 in chronic schizophrenia. Paper read at Amer. Psychiatric Assoc., May, 1955.

4. FREDERKING, W. Ueber die Verwendung von Rauschdrogen (Meskalin und Lysergsaurediaethylamid) in der Psychotherapie. *Psyche*, 1953, 1954, **6**, 342-364.
5. HOCH, P. H., CATTELL, J. P., & PENNES, H. H. Effects of mescaline and lysergic acid (d-LSD-25). *Am. J. Psychiat.*, 1952, **108**, 579-584.
6. ISBELL, H., FRASER, H. F., WIKLER, A., & BELLEVILLE, R. E. Tolerance to Diethylamide of Lysergic Acid. *Federation Proc.*, 1955, Part I, No. 1, 354, and personal communication.
7. LOVELL, R. R. H., OSBORNE, J. A., GOODMAN, H. C., & HUDSON, B. Ascorbic acid, glucose-1-phosphate, and lysergic-acid diethylamide in rheumatoid arthritis. *Lancet*, 1953, **264**, 970-973.
8. SAVAGE, C. Lysergic acid diethylamide (LSD-25): A clinical-psychological study. *Am. J. Psychiat.*, 1952, **108**, 896-900.

133 East 58th Street
New York City 22

PERFORMANCE LEVELS AND TRANSFER EFFECTS IN COMPENSATORY AND FOLLOWING TRACKING AS A FUNCTION OF THE PLANES OF ROTATION OF CONTROL CRANKS*¹

Department of Psychology, University of Rochester

R. F. GREEN, D. GOODENOUGH, B. G. ANDREAS, A. A. GERALL, AND
S. D. S. SPRAGG

A. INTRODUCTION

This report is one of a series of studies in which performance as a function of the relationships between controls and displays of two-dimensional following and compensatory tracking tasks was investigated. Each of these tasks requires that an operator move the controls more or less continuously in response to a continuously changing display. Following tracking requires that the subject keep a follower in contact with a moving target. Compensatory tracking requires that the operator respond to counteract the target movement and to keep the target centered in a null position.

It has been established in previous studies (3, 6, 7) from this laboratory that, other factors being equal, performance will generally be best in a two-dimensional following tracking task when the planes of rotation of the cranks are continuous with the axes of movement of the target follower. Continuity between control and display movements means that the movement of one will be parallel with or "be in line with" the movement of the other. Continuity, as it is used here, is similar to the concept of compatibility which has been used by certain investigators. An example of continuity between controls and display is presented in Figure 1.

It has been considered reasonable to assume that the same generalizations will hold true for compensatory tracking tasks. In two other studies (1, 4), however, it was shown that there is some likelihood that varying the display-control relationships does not affect performance on the compensatory tracking task used (a modified *SAM* Two-Hand Pursuit Test), at least not as

*Received in the Editorial Office on September 12, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹This study was conducted as part of a research program on human motor performance carried out under Contract N6onr-241, T.O. 6 between the University of Rochester and the Special Devices Center, Office of Naval Research, Port Washington, L. I., N. Y., to whom these results have been reported in Technical Report SDC 241-6-22.

much as it does on the following tracking task used (a modified *SJM* Two-Hand Coördination Test). It should be noted that the compensatory task used was not as difficult as the following task. It seems reasonable to expect that control-display relationships become less crucial as a task is made easier.

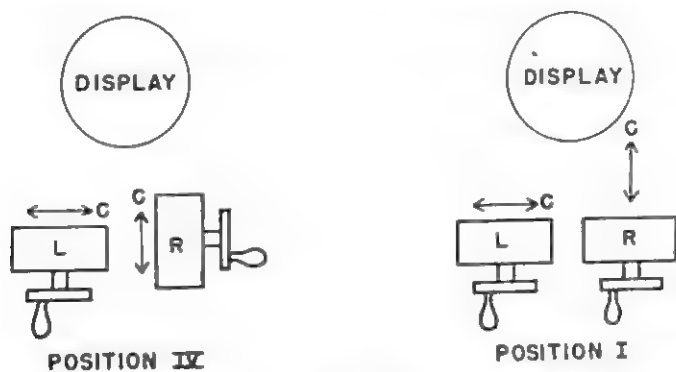


FIGURE 1

TOP VIEW OF THE DISPLAY-CONTROL RELATIONSHIPS

The arrows indicate the dimension of movement of the controlled member of the display which each crank controls. *C* indicates in which direction the controlled member of the display will be driven by a clockwise movement of that control crank.

It is also possible that the fact that the subject's responses are confounded with, and so are partially obscured by, the action of the course generator might result in lessening the importance of display-control relationships in the compensatory tracking task.

One apparent exception to the continuity principle has been obtained. This can be described by considering first a top view of the two display-control arrangements, illustrated in Figure 1. The arrows drawn beside the crank stands indicate the dimension of movement of the controlled member of the display which the crank controls. The letter *C* indicates in which direction the controlled member will be driven by a clockwise movement of the crank.

Display-control movement relations are completely continuous in Position IV but only partially continuous in Position I in which the crank controlling the Y-axis of movement is in a discontinuous relationship. According to the continuity principle, performance on Position IV should be superior to that on Position I. A statistically reliable difference favoring Position IV over Position I was obtained in one study (7) and was not obtained in another study (6) of following tracking. No reliable difference was obtained in one study of compensatory tracking (4). Although the evidence is not conclusive we believe that "true" performance scores obtained by using Position IV are higher than similar scores obtained by using Position I for both kinds

of tracking. But the difference apparently is small enough either to raise a question as to the generality of the effects of continuity on performance or to require that we search for bases for explaining their similarity. Such a basis might be found in the observation that most control situations commonly met in everyday experience involve clockwise movements to drive something away. Lathe handles, vise handles, and right-hand-threaded bolts and screws are examples. Our experiences, then, may correspond more closely to the performance required by Position I than by Position IV. Assuming that skills or abilities are acquired primarily as a result of practice, this would mean that a higher level of previously acquired skill may be available for performing on Position I than on IV at the time of testing. Consequently, scores earned on Position I would be relatively more determined by prior experiences than scores earned on Position IV and, therefore, could be, and in fact apparently are, almost equal to them at the time of testing.

For this explanation to be tenable it would be necessary to show that skills developed on Positions I and IV were not functionally the same. If they were not, then skill developed on Position I would not be expected to transfer completely to subsequent performance on Position IV. Incomplete transfer would mean that it is possible that test performance on Position I could be less than, similar to, or even exceed that for IV, depending on the relative levels of skills which have been developed by a sample of people as a result of their life experience.

B. THE PROBLEM

The immediate problem investigated was to determine transfer effects between Positions IV and I in both compensatory and following tracking tasks. The more general problems were to obtain information on transfer effects between a continuous and a partially continuous display-control movement relationship and to compare these effects in the two kinds of tracking situations.

C. SUBJECTS

Eighty-two right-handed basic trainees at a military installation served as Ss. These Ss were divided at random into four groups. Two groups of 20 each were tested on the compensatory tracking task, and two groups of 21 each on the following tracking task.

D. PROCEDURE

Parallel procedures were followed for each kind of tracking. All Ss were given 16 one-minute trials. There was a 15-second interval between trials

for groups performing the compensatory task and 30-second intervals between trials on the following task. The difference in interval between trials was the result of more time being needed to record the multiple scores secured for the following tracking task in contrast to the compensatory task.

A simple *AB, BA*, transfer design was employed. One of the two groups assigned to the same tracking task was given 8 trials on Position I followed immediately by 8 trials on Position IV. The other group was given 8 trials on Position IV followed immediately by 8 trials on Position I. Approximately one minute more rest occurred between the 8th and 9th trials than between other trials, during which time the position of the control was changed. No practice with the crank in its new location was permitted before the start of Trial 9.

E. APPARATUS

1. *Compensatory Tracking Apparatus*

A modified version of the *SAM* Two-Hand Pursuit Test was used. A detailed description of the original test (CM810B) may be found in AAF Report No. 4 (5). The modification of the Pursuit Test controls has been described in a previous report from this laboratory (4). The modification pertinent to this study consisted of mounting the control cranks on pivoting stands. A coupled reversible synchro drive system was then installed between the handles and the test drive mechanism. By means of an enclosed mirror system, arranged as a horizontal periscope, the display was brought to the side of the apparatus close to the *S* rather than remaining in the center of the apparatus as in the original test. This change made it possible for *S* to view the display with less effort than on the original test, and made the *Ss'* operating posture more nearly comparable to that on the following tracking task.

The task was to keep a 1/5" diameter silver disk (the target) centered under a set of crosshairs. The target, which was viewed through a 3-inch diameter aperture, was moved out of alignment with the crosshairs by a motor which rotated the black turntable into which it was imbedded. *S* attempted to counteract the perceived target movement by rotating either or both of the two cranks, which were 19" apart. Superimposing *S*'s responses upon the target movement in this way made the target movement appear to be irregular. Performance was scored as the time per trial the target was kept under the crosshairs. A Standard Electric Timer was used.

2. *Following Tracking Apparatus*

A modified *SAM* Two-Hand Coördination test was used. This test (CM101B) is also described in detail in AAF Report No. 4 (5). The modifications of the controls, which were exactly the same as those for the compensatory test, have been described in detail elsewhere (6). The control cranks were identical in size and location for the two tests. Ss attempted to keep a target follower on a 5'8" diameter brass target. The target was moved through an irregular circular course by the rotation of a horizontal 7-inch diameter turntable and the action of a set of fixed and moving course cams.

The scoring of following tracking performance was increased considerably for this study by an arrangement which produced a total of 10 scores for each trial. The usual time on target score was obtained and this score was associated with a counter score of the number of times the subject got on (or lost contact with) the target. In addition to these two scores, an arrangement was worked out which made it possible to obtain eight different error scores. An error was defined as turning the crank in a direction which would move the follower away from instead of toward the target. A separate clock was provided for recording the total time per trial of such errors in each of the four directions; also a separate counter recorded the total number of times that errors occurred in each direction. One lack in this error scoring system was that it took no account of rate errors. But it seemed desirable to obtain the scores which were chosen, in order to determine whether it would be fruitful to have a more detailed analysis of performance than is afforded by time on target scores.

F. RESULTS

1. *Compensatory Tracking*

The mean time on target scores for each group on each trial are presented graphically in Figure 2. Inspection of this figure shows that mean performance was higher on Position IV than on Position I during the original learning trials. This difference was statistically significant (t -ratio between grand means was 2.38 with 38 df).

From a comparison of Group *B*'s transfer trials with Group *A*'s original learning trials, it is apparent that transfer effects from Position IV to Position I were positive and persistent for all 8 transfer trials. This difference was also significant (t -ratio between grand means was 4.38 with 38 df). The data also permit a comparison of the transfer trials of Group *A* with the

corresponding original learning trials of Group *B*. This shows that transfer effects from Position I to Position IV were initially positive. The difference between the first trial means of the two groups (Group *B*, original learning vs. Group *A*, transfer) is significant ($t = 2.47$, 38 *df*), but the differences

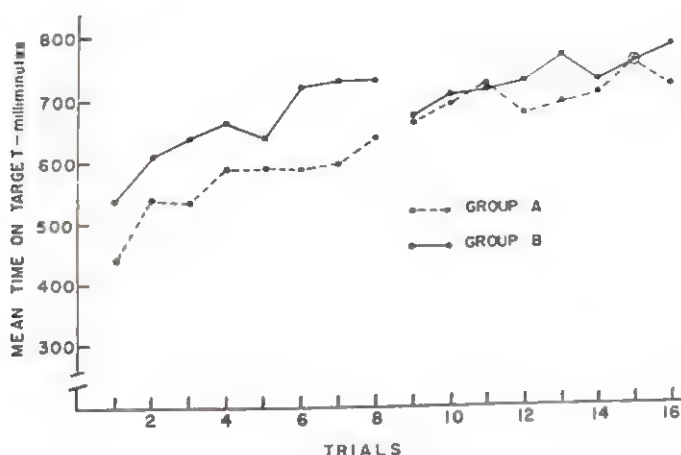


FIGURE 2
COMPENSATORY TRACKING TASK

Mean time on target per trial for the eight original learning trials and the eight transfer trials for Group *A* (Position I, then Position IV), and Group *B* (Position IV, then Position I).

between corresponding trials of the two groups becomes progressively less and after the third trial there is no evidence of a persistence of transfer effect.

It should be noted in Figure 2 that performance on the two positions was not reliably different after transfer. Practice on Position IV appeared to have transferred completely to subsequent performance on Position I, but practice on Position I did not transfer completely to subsequent performance on Position IV.

2. Following Tracking

The means of the time on target scores of each group for each trial are shown graphically in Figure 3. It is evident from this figure that there was no reliable difference between the time on target scores of the two groups during the 8 original learning trials. There was, however, a significant difference ($t = 2.023$, 40 *df*) between the grand means of the eight transfer trials of the two groups. This difference favored the group performing on Position IV.

Both groups exhibited strong positive transfer effects. There is evidence,

however, to suggest that transfer in Group *D* from Position IV to Position I was not as complete as transfer in Group *C*. Inspection of Figure 3 shows that Group *C*'s performance on Position IV would closely approximate an extrapolation of the original learning trials on that position. However, the

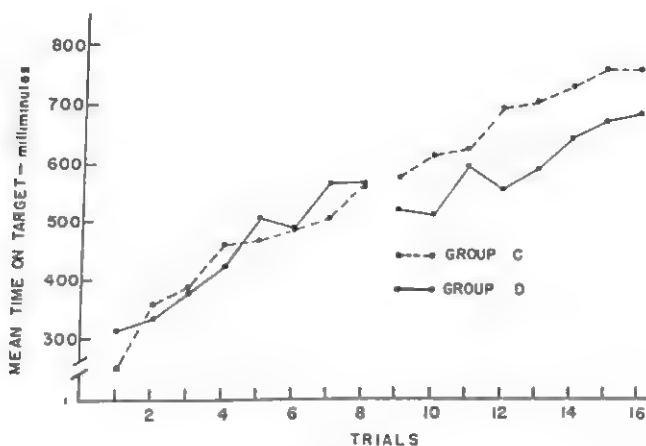


FIGURE 3
FOLLOWING TRACKING TASK

Mean time on target per trial for the eight original learning trials and the eight transfer trials for Group *C* (Position I, then Position IV), and Group *D* (Position IV, then Position I).

transfer performance of Group *D* on Position I is somewhat below any extrapolation that might be made from the original learning curve. Further support for the finding that transfer was different in the two groups was the fact that their mean performances were reliably different on the transfer trials, as noted above, while their original learning scores had been about equal. A more direct test of the differential transfer effects, $(\text{Position IV}_T - \text{Position IV}_{OL}) - (\text{Position I}_T - \text{Position I}_{OL})$, did not reach statistical significance.

Nine additional measures of following tracking performance were obtained (see Part 2 of apparatus section, above). The mean counter scores for frequency of target contacts were equivalent for the two groups, before and after transfer. Since the associated time on target scores differed, this means that during the transfer trials Ss in Group *D* got on target as frequently as the Ss in Group *C* did but did not stay on target as long.

The total time per trial that the Ss spent in cranking away from the on-target position was analyzed for each of the four primary directions and also for the combinations *left-right* and *toward-away*. Differences between *left*

and *right* error means were not significant, either for groups or conditions. Differences between *toward* and *away* error means were found to be significant in certain cases. Group *C* had a larger *toward* duration of error mean ($t = 3.36$) during their original learning trials on Position I. This difference persisted but to a lesser extent during their subsequent transfer trials on Position IV. However Group *D* had no reliable differences in mean error duration during either original or transfer trials.

Within each group, there were reliably higher combined *toward-away* error means than combined *right-left* error means (Group *C*, $t = 9.13$, Group *D*, $t = 8.19$) during the original learning trials. This difference persisted throughout the transfer trials. There was, then, a reliable overall tendency for each group to have a higher *toward-away* duration of error mean than *right-left* error mean despite the fact that the *toward-away* cranking was done by the right hand.

There were counter scores associated with each of the four clock scores discussed above which measured the frequency of cranking in incorrect directions. These results will not be presented in detail. In the main the counter score results were similar to the clock results. In those instances in which the two were not in agreement, the counter score differences were usually less significant than the clock score differences.

Estimates of reliability and of intercorrelations of the 10 performance scores were obtained. The correlational evidence gave further support to the statement made above that the counter score results in general were similar to the clock results and hence did not add significantly to our understanding of two-hand following tracking performance.

G. DISCUSSION

The compensatory tracking results appeared to differ in certain important respects from the following tracking results. In the former, performance on Position IV was reliably better than on Position I during the original learning; in the latter no reliable difference between positions was obtained. This particular difference was unexpected. Position IV was the one having complete continuity between display and controls and was therefore expected to result in better performance in both kinds of tracking. So far in our series of tracking experiments in studies involving other variables in addition to this one, there have been six independent tests of the difference between Positions I and IV, two of which have been significant at the 5 per cent level of confidence. The probability of obtaining two significant results out of six tests by chance is .04 (8). Two of these tests involved compensatory and

four involved following tracking. One significant result occurred on each. So even though the combined probability is significant, it is not significant for either kind of tracking taken separately. It is likely that eventually Position IV will be proved to result in statistically better performance than Position I, but the difference may not be enough to be of practical significance.

Compensatory and following tracking also appeared to differ in transfer effects. In compensatory tracking, based on extrapolation of the original learning curves, the effects of transferring from Position IV to I were positive and probably persistent and complete. Transfer effects from I to IV were positive but neither complete nor persistent. That is, transfer from the harder to the easier task was not complete. Stated still another way, original practice on Position IV (the easier task) was equivalent to the same amount of practice on Position I (the harder task) while the converse was not true. This result is contrary to the principle cited by Gibbs (2), to the effect that there is generally greater transfer from a difficult to an easier version of a task than there is from the easy to the difficult. The finding here may call for an extension of Gibbs' suggestion that the differential transfer effects based on differing task difficulties may arise from different causes. It may be that some methods of varying task difficulty do not conform to the principle that difficult-to-easy leads to greater positive transfer. The difficulty concept, as it relates to transfer effects, may have to be replaced by a number of separate concepts—each related to task difficulty—and each may bear any one of several relations to transfer.

While the transfer results from following tracking were different from compensatory tracking effects, they do not strictly conform to usual expectations either. Here, the two positions resulted in almost identical levels of performance during original learning. But transfer from Position IV to I, while positive and persistent, apparently was incomplete; and transfer from Position I to IV was positive, persistent, and complete. A conclusion reached by Gibbs (2) suggested that if the two tasks were equally difficult transfer effects should have been equivalent.

There is another way of analyzing the results. If we ask what happens to performance scores when operators are switched from one of the positions used to the other, the changes in each group's performance levels appear quite similar for both kinds of tracking. When the groups transfer from Position IV to Position I, their performance, at least on the first few trials after transfer, shows a decrement. That is, their performance on Position I is lower than it had been at the end of training on IV. When the groups

transferred from Position I to IV no such decrement occurred. That is, for both kinds of tracking subsequent performance on Position IV continued along the same function that had been shown on Position I. In both cases, then, when transfer was from a completely continuous to a partially continuous display-control relationship a decrement in the performance of the groups concerned was observed. When transfer was from a partially continuous to the completely continuous display-control relationship no such decrement was obtained. These observations seem to support the idea that Position IV is intrinsically easier than Position I in that here the change from IV to I leads to a decrement in performance while the change from I to IV does not lead to a drop in absolute level of performance of the group involved.

From the error scores, it was shown that the average duration of errors made in the *toward-away* (Y) axis was larger than in the *right-left* (X) axis. This difference occurred on both positions. Whether it was due to intrinsic differences between hands or between axes was not definite since effects of hands and axes were confounded in the experimental design used. But since the greater average duration of errors was made in the axis controlled by the dominant hand, it might be expected that, at least for this device, the *toward-away* (Y) axis of movement in a horizontal display will prove to be intrinsically more difficult to track than the *right-left* (X) axis.

Related to the above result was the observation that transferring to Position I appeared to result in somewhat greater number of errors made by each hand—but particularly the right hand, which operated the control that was changed—than transferring to Position IV. There was, however, no severe tendency for the average duration of errors made with either hand to increase as a result of moving the position of a control. There also appeared to be a somewhat greater tendency for the mean duration of the *toward* errors to be larger than that of *away* errors for Position I. This tendency was not evident for Position IV. It would appear, then, that the main difference between Positions I and IV might be that Position I is more conducive to subjects' spending more time cranking in the wrong direction in the Y-axis—particularly errors toward the subject.

Transfer effects between Positions I and IV for both following and compensatory tracking were important in this study principally because of their implications pertaining to the continuity principle. Although the overall results again were somewhat equivocal they do at least partly support the principle that complete continuity between planes of control and display movements will be associated with better performance than if such continuity

is not completely present. Transfer from a continuous to a partly discontinuous arrangement results in a larger performance decrement than transfer from partial discontinuity to continuity.

H. SUMMARY

The problem was to investigate performance levels and transfer effects in both compensatory and following tracking as a function of two arrangements of display-control movement relationships.

Eighty-two *Ss* from a military population were divided into four groups, two for compensatory and two for following tracking. A simple *AB, BA* transfer design was used for each kind of tracking.

Modified versions of the *SAM* Two-Hand Coördination Test and *SAM* Two-Hand Pursuit Test were used for the following and compensatory tasks, respectively. The left hand crank control rotated in a plane parallel to the frontal body plane of the *S* for all groups. For two groups the right hand crank also rotated in the frontal plane (called Position I) and then after eight trials the crank mount was turned 90° to the right so that the crank then rotated in a plane perpendicular to the frontal body plane (called Position IV). The other two groups started with the right crank in the latter position (IV) and after eight trials it was moved to the former position (I). The principle that continuity between planes of movement of display and controls will result in optimum performance would lead to the prediction that in both kinds of tracking Position IV should yield the better performance during original learning.

As has been the case in some of our previous studies, the results pertaining to initial performance levels on the two positions were somewhat equivocal. A statistically significant difference in favor of the continuous display-control movement relationship occurred in compensatory tracking, but no such difference occurred in following tracking. Evidence was presented, based on the combined results of six independent tests, showing that the hypothesis that Position IV (which has complete continuity between control and display movements) will lead to better performance than Position I (in which the right control movement is discontinuous with the display movement) can still be considered tenable.

Analysis of transfer effects showed that in compensatory tracking transferring from the continuous (easier) relationship to the partially continuous (harder) relationship yielded better transfer than transferring from the harder to the easier relationship, while the reverse relations held for following tracking.

However, analysis of the performance change for each particular group when a change in task was imposed, showed similar effects for both kinds of tracking. In each case a change from continuous to partly discontinuous display-control relationships produced a decrement in performance, while the reverse change produced no such decrement.

In addition to time on target, nine other performance scores were obtained. Eight of these can be called error scores when error measures include the average length of time and the number of times a subject turned his controls in the wrong direction. More time was spent cranking in the wrong direction in the *toward-away* (or *Y-axis*) direction than in the *right-left* (or *X-axis*) direction. This was true whether the *Y-axis* control was in the continuous relation (Position IV) or not (Position I).

REFERENCES

1. ANDREAS, B. G., GREEN, R. F., & SPRAGG, S. D. S. Transfer effects between performance on a following tracking task (modified *SAM* Two-Hand Coördination Test) and a compensatory tracking task (modified *SAM* Two-Hand Pursuit Test). *J. of Psychol.*, 1954, **37**, 173-183.
2. GIBBS, C. B. Transfer of training and skill assumptions in tracking tasks. *Quart. J. Exp. Psychol.*, 1951, **3**, 99-110.
3. GREEN, R. F., ANDREAS, B. G., NORRIS, E. B., & SPRAGG, S. D. S. Performance on a following tracking task (the *SAM* Two-Hand Coördination Test) as a function of the continuity of the plane and direction of movement of the control crank and target follower. *J. of Psychol.*, 1955, **40**, 403-410.
4. GREEN, R. F., NORRIS, E. B., & SPRAGG, S. D. S. Compensatory tracking performance (modified *SAM* Two-Hand Pursuit Test) as a function of the directions and planes of movement of the control cranks relative to movement of the target. *J. of Psychol.*, 1955, **40**, 411-420.
5. MELTON, A. W., (Ed.). Apparatus Tests. Report No. 4, U. S. Army Air Forces Aviation Psychology Reports. Washington, D. C.: U. S. Government Printing Office, 1947.
6. NORRIS, E. B., & SPRAGG, S. D. S. Performance on a following tracking task (modified *SAM* Two-Hand Coördination Test) as a function of the planes of operation of the controls. *J. of Psychol.*, 1953, **35**, 107-117.
7. ———. Performance on a following tracking task (modified *SAM* Two-Hand Coördination Test) as a function of the relations between direction of rotation of controls and direction of movement of display. *J. of Psychol.*, 1953, **35**, 119-129.
8. WILKINSON, B. A. A statistical consideration in psychological research. *Psychol. Bull.*, 1951, **48**, 156-158.

Department of Psychology
University of Rochester
Rochester 3, New York

AN EXPERIMENTAL METHOD TO MEASURE INTELLECTUAL FUNCTIONING WITH VERBAL AND MOTOR FACTORS MINIMAL (1)*

Department of Psychology, New York University

LEONARD BLANK AND MOSS L. RAWN

A. PURPOSE

The purpose of this study was to devise a method of objectively measuring perceptual and organizational processes relevant to intellectual functioning. These processes were defined operationally and the Method constructed according to the criteria thus imposed. An effort was made to keep verbal, motor, and cultural factors minimal.

The following problems were investigated: (a) Is intelligence being measured? (b) Does this intellectual functioning increase with age? (c) Is this functioning relatively free of verbal and motor factors? (d) Do these intellectual factors consist largely of the operationally defined perceptual and organizational processes?

B. HISTORY

Intelligence is generally understood to be composed of more than any one or group of abilities such as verbal or motor (4, pp. 215-223), (6, chs. 4-6), (7, chs. 10-12), (9, chs. 4-5), (11, ch. 5). Most measures of intelligence, however, include only one of these abilities or combine both. Perceptual and organizational processes are relatively neglected (5). Motor handicapped are tested primarily with verbal material while verbally handicapped are given motor tasks (5). In both cases, a distorted picture of intelligence is presented. There is a lack of objective measurements for the *both* verbally and motor handicapped of which the Cerebral-Palsied make up a large part (2, 3, 8, 10). Non-verbal and "culture-free" tests of intelligence include pencil manipulation, speed factors, and abstract material (4, 6).

The Experimental Method was designed to measure objectively perceptual and organizational processes and thus help to provide a more comprehensive representation of intelligence with special application to handicapped and non-literate persons. It was constructed to present a situation where verbal, motor, cultural, speed, and abstract features were minimal.

*Received in the Editorial Office on September 15, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

C. THE EXPERIMENTAL METHOD

The Method consisted of multiple-choice visual discriminations. The subjects assembled a male figure from among original pieces and alternates of subtle, moderate, and gross distortion but all of interchangeable fit. A practice figure was first demonstrated in pantomime and then 11 sets of four pieces each were presented in prescribed order. The pieces were composed of photostats of life-like drawings glued to three-fold plywood (1, p. 181 for photograph of presentation of Method.)

Subjects indicated their choice by touching a piece in some way. No verbal or motor responses were elicited and there was no time limit. Perceptual processes were defined as immediate and progressive discriminations of the various body parts, one from another and from alternate pieces of the same part. Organizational processes were the combining of these percepts into larger and larger units as well as the correction of these percepts.

D. POPULATION

The test population consisted of 12-, 14-, and 16-year-old males of dull, average, and bright intelligence. Forty-nine students were sampled at a Junior and a Senior High School in New York City; 29 Mental Defectives at a state institution; 67 Delinquents at a State Training School; and 48 subjects at out-patient cerebral palsy clinics. The total sample was 193 subjects. The Method was administered individually to each subject and Form I of the Wechsler Adult Intelligence Scale to the students and Delinquents. The Defectives had been administered Stanford-Binets. The intelligence of the Cerebral-Palsied was evaluated by psychologists who worked with them.

E. RESULTS

The pantomime instructions were grasped readily by all subjects, including the Mental Defectives. The administration proved efficient with a maximum time of 20 minutes and an average time of 12 minutes.

Tests for Homogeneity of Variance demonstrated that differences between mean Method score were not attributable to variability within the samples. Analysis of Variance revealed that the variables of age and intelligence did not contribute significantly to the variance differences when over-all groups described by age and intelligence were considered. These two variables were indicated as being relatively independent of each other.

In Table 1 are presented the mean experimental score and standard deviation for each sample population. Although differences of mean Method score between age and intelligence sub-groups were not statistically signifi-

cant, trend differences were established. This was demonstrated by the Sign Test for Significance (12, pp. 431-432). Thus, in the comparison of mean scores where intelligence was varied, 30 out of a total of 36 differences were in the predicted direction—brighter subjects scored higher than duller. The Sign Test demonstrated that the probability of this occurring due to chance is less than one in a thousand. Mean experimental score thereby is significantly and positively associated with intelligence classification. (It will be seen later that this finding is substantiated by correlations of the Method with Wechsler Intelligence Scores.)

TABLE 1
EXPERIMENTAL METHOD MEANS AND STANDARD DEVIATIONS

Students	Number	Mean	Standard Deviation
Students	49	36.06	4.02
Delinquents	67	34.91	4.46
Cerebral-Palsied	48	32.40	4.27
Mental Defectives	29	32.28	4.62

When both intelligence and age comparisons are jointly considered, 50 out of 67 differences are in the hypothesized direction. The probability that this is due to chance alone is much less than one in a thousand.

Both intelligence and age were, therefore, positively associated with Method score.

F. INTER-SAMPLE COMPARISONS

The following sample order emerged in almost every analysis: Normals scored highest; Delinquents, second; Cerebral-Palsied, third; and Mental Defectives, lowest.

This hierarchical sample order may be noted in Tables 1, 2, and 3. In the latter two tables, the general increase of Method score with intelligence and with age may also be noted.

All four sample populations as wholes are compared for statistical significance of mean Method scores in Table 4. Five out of 6 of these differences are statistically significant. Only Cerebral-Palsied are not statistically differentiated from Defectives although they do score higher. The sample order was again manifested as it was in every analysis of large group mean score comparisons.

The Sign Test of Significance was applied to the inter-sample comparisons of mean Method scores. For the comparisons of sub-group means by age and by intelligence, 24 out of 29 lie in the hypothesized direction. This is significant beyond the one per cent level of confidence. When the larger

categories of all subjects within each age, intelligence, or sample category are considered, 23 out of 24 mean differences lie in the predicted direction. Such a result would occur less than one in a thousand times by chance alone. If these comparisons are considered together, then 47 out of 53 are in the expected direction. Again this would occur less than one in a thousand times by chance.

TABLE 2

MEAN METHOD SCORE FOR EACH AGE AND SAMPLE IRRESPECTIVE OF INTELLIGENCE

Sample	12 Yr. olds	14 Yrs.	16 Yrs.	Means
Students	35.69	36.61	35.80	36.06
Delinquents		34.40	35.06	34.91
Cerebral-Palsied	31.81	33.47	32.00	32.40
Mental defectives	31.14	32.08	33.30	32.28
All samples	33.28	34.37	34.44	34.18

TABLE 3

MEAN METHOD SCORE FOR EACH INTELLIGENCE AND SAMPLE IRRESPECTIVE OF AGE

Sample	Defectives	Dull	Average	Bright	Means
Students		34.60	35.82	37.59	36.06
Delinquents		34.28	35.42	36.43	34.91
Cerebral-Palsied		31.82	32.11	33.67	32.40
Defectives	32.28				32.28
All samples	32.28	33.74	34.48	36.06	34.18

TABLE 4

t-SCORE COMPARISONS BETWEEN ALL THE SAMPLES

Group	t	P	Samples With Higher Mean
N.CP	4.31	.001*	N
N.DL	2.08	.02*	N
N.DF	3.60	.001*	N
CP.DL	3.02	.001*	DL
CP.DF	.11	.90+	CP
DL.DF	2.55	.01-.02*	DL

*Significant at the .05 or less level of confidence.

G. CORRELATIONS OF THE METHOD WITH WECHSLER SCORES

Eta correlations (12, pp. 276-278) between Method and Wechsler Adult Intelligence Test scores were computed for the student and Delinquent samples. The Eta coefficient for the students was .726, which was both significantly different from chance and from linearity at the one per cent level of confidence. That is, when predicting Method from Full Scale Wechsler score, the Eta coefficient of .73 best describes this relationship (whereas the product-moment r , because of curvilinearity, would tend to

underestimate this relationship.) The *Etas* for the Delinquents were insignificantly different from both chance and linearity so that *r* best describes the relationship with Wechsler Scores.

TABLE 5
CORRELATIONS OF THE EXPERIMENTAL METHOD WITH WECHSLER SCALES AND SUBTESTS
FOR THE STUDENT AND DELINQUENT SAMPLES

Scale or subtest		Students		Delinquents	
		<i>IQ</i> or weighted score	<i>r</i>	<i>IQ</i> or wt. score	<i>r</i>
Verbal		101.29	.354*	87.38	.345*
Performance		103.20	.341*	99.73	.190
Full		102.92	.371*	92.53	.380*
Information		9.12	.321*	7.24	.275
Comprehension	Conventional				
	Judgment	10.18	.243	8.56	.249
Digit Span	Attention	7.61	.035	6.76	.117
Arithmetic	Concentration	7.76	.381*	6.04	.342*
Similarities	Verbal Concepts	9.53	.462*	7.56	.357*
Vocabulary	Organization of Memories and Concept Formation	8.65	.495*	7.00	.391*
Picture Arrange- ments	Anticipations—Visuo- motor Grasp of Social Situations	10.12	.058	10.20	.165
Picture Completion	Recognition and Grasp of Whole Situations	10.06	.375*	10.11	.365*
Block Design	Visuo-motor Coordina- tion and Organization	9.98	.313*	9.36	.131
Object Assembly	Visuo-motor Coordina- tion and Anticipation	9.39	.364*	11.13	-.099
Digit Symbol	Visuo-motor Coordina- tion—Concentration and Learning	9.02	.162	7.53	.150
N for Students equals 49					
N for Delinquents equals 45					

*Significant at .05 level of confidence.

Table 5 presents product-moment correlations of the Method with the Full, Verbal, and Performance Scales and with the weighted scores of the various subtests of the Wechsler for the student and Delinquent samples.

Correlations¹ for both samples were highest for Full *IQ*, second highest

¹The differences between the correlations of the Wechsler scores with the Experimental Method that are statistically significant are as follows: For the Students, the Vocabulary and Similarities correlations are both significantly greater than the Digit Span and Picture Arrangement coefficients; For the Delinquents, the Full, Verbal, Vocabulary, Similarities, Picture Completion, and Arithmetic correlations are all significantly higher than the Object Assembly coefficient; the only correlation significantly different between the Students and the Delinquents is the Object Assembly coefficient in favor of the Students.

for Verbal, and lowest for Performance *IQ*. These correlations were statistically significant save for Delinquent Performance *IQ* with Method. The highest correlations for both samples were for Vocabulary and Similarities and for Picture Completion in the Performance Scale. Despite as high or higher scores on Wechsler subtests involving visuo-motor coördination and speed factors, the Delinquents did not obtain statistically significant correlations with these subtests with the Method while the Normals did. This finding suggested that the Method is not measuring these factors but probably more of "g" and/or perceptual-organizational processes.

H. RELIABILITY

There were three problems of reliability: (a) How internally consistent was the Method? (b) What was the effect on mean scores of administration by different examiners, or inter-tester reliability? (c) Is Method score consistent from one administration to another?

In Table 6, the split-half corrected reliability coefficient between mean Method scores for odd and even pieces for the student sample is .55 and is statistically significant while the difference between these means is insignificant: the Method is internally consistent for this sample. The difference of mean Method scores between investigators is insignificant for all samples:

TABLE 6
SPLIT-HALF, INTER-TESTER, AND TEST-RETEST INDICES OF RELIABILITY

Source	Statistic	Coef.	<i>t</i>	<i>P</i>
Odd and even pieces	Split-Half <i>r</i>	.382	2.67	.004*
Students	Spearman-Brown <i>r</i>	.553	3.87	.001*
Delinquents	Inter-Tester <i>t</i>		.091	.90
Cerebral-Palsied	Inter-Tester <i>t</i>		.596	.50
Mental Defectives	Inter-Tester <i>t</i>		.465	.60
Students	Inter-Tester <i>t</i>		.20	.80
(N=14)	Test-Retest	.323		.40-.50
	Test	Re-test	<i>t</i>	<i>P</i>
Mean	36.14	37.79	1.02	.40-.50
Standard Deviation	4.48	3.79		

the Method demonstrates inter-tester reliability. The test-retest coefficient of reliability for the students is .323 and is statistically insignificant. This may represent a serious limitation for the Method. However, the original mean score is statistically insignificantly different from the retest mean score. A possibility of learning may also enter the situation since the retest mean is higher and the index of variability lower. The one-month interval between testing may have been too low. Further, the small *N* of 14 may also be somewhat responsible for the low test-retest coefficient of reliability.

I. DISCUSSION

The preceding results may be interpreted as positively answering the first two problems under investigation, i.e., the Method discriminated among different intelligence groups with the same age and among different age groups with the same intelligence. Statistical significant differences may not have been the rule but trends of mean differences for the age and intelligence variables were established.

The hierarchical sample order that emerged appeared to be in accord with a functional concept of intelligence. Thus, students who were relatively undisturbed by physical and emotional impairment, compared to subjects so afflicted of the same age and intelligence in the other samples, fared clearly superior on the Method. Delinquents who, by definition, suffered emotional difficulty, fared second best. The Cerebral-Palsied, who were afflicted with brain damage and possibly with concomitant emotional difficulty, scored third. The Mental Defectives were lowest in intelligence and lowest on the Method.

The third problem under investigation also seems to have been answered positively. The Method has been shown to measure generalized intellectual functioning while at the same time it manifested less relationship with verbal and motor factors. Positive answers to the first two problems were part of this substantiation. The discrimination between Normals and Defectives on the Method in favor of the former in every analysis and always statistically significant, when considering the samples as wholes, was added proof that intelligence was being measured. The statistically significant correlations with the Full Wechsler *IQ* and significant correlations of lesser magnitude with Verbal and Performance *IQ*'s, for both students and Delinquents, provided indices of external validity in this respect as well.

In light of the .73 (Eta coefficient) correlation of the Method with The Wechsler Full Scale, the Method measures that type of ability being measured by the Wechsler scales as well as or better than the individual subtests of these scales. Yet for both the students and the Delinquents, the Method correlates higher with the Full Scale than with either the Verbal or Performance Scales. It may be concluded that the Method is measuring intelligence while at the same time manifesting less relationship with verbal and motor factors.

The last problem posed also appears to have been resolved positively: the Method correlates as well or better with the Wechsler Full Scale than any of the latter's subtests; it correlates well with subtests high in "g" and in perceptuo-organizational processes and low in visuo-motor coordination and

is latent in the mind of one who has learned to write; the second is active, as when the scribe evokes from the latent state this knowledge of writing which he desires to put into practice; the third is the degree of intelligence actually involved in the operation of writing, where the knowledge now quickened into activity guides and directs the act. The external faculty is the "Agent Intellect" (*'aql fa''al*) which proceeds from God by way of emanation and which, though acting on the faculties in the body, is independent of the body, as its knowledge is not based upon perceptions obtained through the senses.

These four aspects of intelligence become somewhat clearer as described in the following account of another philosopher, namely, al-Farabi (died 950 A.D.) from O'Leary (6):

Following al-Kindi, al-Farabi speaks of four faculties or parts of the soul: the potential or latent intelligence, intelligence in action, acquired intelligence, and the agent intelligence. . . . The agent intelligence or *'aql fa''al* is the external power, the emanation from God which is able to awaken the latent power in man and arouse it to activity, and the acquired intelligence or *'aql mustafad* is the intelligence aroused to activity and developed under the inspiration of the agent intelligence. Thus the intelligence in action is related to the potential intellect as form is to matter, but the agent intelligence enters from outside, and by its operation the intelligence receives new powers, so that its highest activity is acquired.

Students of the modern mental test movement may recognize here certain distinctions which have played a part in recent psychological theory. Thus intelligence tests were once thought to be indices of native capacity alone, i.e., potential or latent intelligence. But further experience with tests led to the belief that they were better described as tests of actual or effective intelligence (the intelligence in action of al-Farabi). As for the third aspect—the acquired intelligence of the Arabian—we may compare it with the view of one of our most thoughtful authorities in this field. I mean the definition of intelligence once given by V. A. C. Henmon, when he said that "intelligence is the capacity for knowledge and the knowledge possessed."

Among other problems discussed by the Arabian writers were those of the relation of mind to abstract ideas. The dispute over universals, so important in later western philosophy, is anticipated in al-Farabi's position. According to the opinion of de Boer, Farabi has envisaged the relationship correctly when he says that existence is a grammatical or logical relationship, but not a category of reality which tells anything about things. The existence of a thing is nothing but the actual thing itself. Being has no reality aside from the things that be.

Another great name among Arabians is that of Ibn Sina or Avicenna (980-1037). He too taught that "the generic belongs only to the realm of thought" and that "abstract ideas have no objective existence, although they may be used as real in logic" (quoted from O'Leary, *op. cit.*, p. 75). Further features of Ibn Sina's psychology are thus given by O'Leary (6).

The faculties present in the soul may be divided into two classes, the faculties of perception and the faculties of action. The faculties of perception are partly external and partly internal: of these the external faculties exist in the body wherein the soul dwells and are the eight senses, sight, hearing, taste, smell, perception of heat and cold, perception of dry and moist, perception of resistance as by hard and soft, and perception of rough and smooth. By means of these senses the form of the external object is reproduced in the soul of the percipient. There are four internal faculties of perception: (i) *al-musawwira*, "the formative," whereby the soul perceives the object without the aid of the senses as by an act of imagination; (ii) *al-mufakkira*, "the cogitative," by which the soul perceiving a number of qualities associated together abstracts one or more of them from the others with which they are associated, or groups together those which are not seen as connected; this is the faculty of abstraction which is employed in forming general ideas; (iii) *al-wahm*, or "opinion," by means of which a general conclusion is drawn from a number of ideas grouped together; and (iv) *al-hafiza* or *al-zakira*, "memory," which preserves and records the judgments formed (p. 176).

Algazel (1058-1109 or 1111), surnamed "Islam's convincing Proof," was a thinker whose position among Moslems has been compared to that of St. Thomas Aquinas among Christians. The following paragraphs give some of his arguments for the existence of higher thought process better than any we now know. (Translated from Miguel Asin Palacios: *Algazel* (1). The philosopher is speaking to himself:

"From whence comes the confidence which I place in the perceptions of the senses? Among these the sight is apparently the faculty most assured in its perceptions. Well then; if I look at a shadow I see it fixed, immobile, and I judge by this that it lacks movement; but, an hour later, I shall know, by experience and observation, that the shadow has moved, for its movement does not verify itself suddenly and all at once, but by degrees and little by little, so that it is never still. In the same way vision regards the stars and sees them so small that their size is no greater than a peseta; but afterwards come the geometrical demonstrations to prove that they are larger than the earth in magnitude."

"All these phenomena and others like them are testimonies for the senses, which give, concerning them, definite judgments, which the understanding discredits and declares false, without their being able to repel or to refute this accusation."

"Tell me then: if this confidence which you have put in the perceptions of the senses should disappear, perhaps no guarantee of certainty exists more than in the first intelligibles, that is to say, in the first principles, such as these: twelve is more than three; negation and affirmation cannot join together in one single thing; the same object is not simultaneously temporary and eternal, being and not being, necessary and impossible."

"But the perceptions of the senses exclaimed: and who guarantees to you that your trust in the first principles is not identical with the trust which previously you had in our perceptions? Because, when you then were sure of us, there came the understanding and belied us; so that, if you had not had the judgment of the understanding, you would have continued still to give credit to our testimony. Well then, perhaps there exists, beyond the intellectual perception, some other judge or criterion which, if it should manifest itself to us, would belie reason in its judgment, as reason manifested itself when it came to belie the judgment of the senses. And, the fact that the said criterion does not show itself to us is no proof of the impossibility of its existence."

"My spirit remained perplexed for some time, without knowing how to answer this objection; but, far from emerging from these doubts, they confirmed themselves the more, as I reflected upon dreams."

"Do you not perchance notice, said I to myself, how in a dream you believe that you see things, and imagine definite situations to which you assign reality, giving absolute credence, without a single doubt assailing you about them, while you are sleeping? And, nevertheless, when you awake, you know that evidently all those things which you imagined and to which you gave credence, lack foundation and value. Then, who assures you that all which you believe that you know in the waking state, with the senses or with the reason, has objective reality? Certainly it is all true in relation to your state; but it can very well happen that you come to experience in another state that it should be with respect to the awakened state that which this is with respect to dreaming; hence it will come to pass that that which now you call waking is dreaming, and you will know with certainty that all which you judged with your reason are chimeras of phantasy without any reality."

"Who knows if this superior state be not the same which the sufis name their ecstasy? . . ."

This was indeed the conclusion to which Algazel came. He thought that there must be some higher state of intuition. In order to attain it he surrendered a brilliant career, as a teacher in Bagdad, and leaving his wealth for the support and education of his family, he entered upon a new career of study and of ascetic practices. His mystic philosophy was in no wise opposed to the Quran. He was a defender of orthodoxy. He held that God is the Creator of all things, even of the free acts of men.

Algazel's famous book, *Destruction of the Philosophers*, was answered by another famous treatise, *Destruction of the Destruction*. The latter was by Ibn-Rushd or Averroes of Cordoba (1126-1198). Averroes earned the name of the Great Commentator, in recognition of his service in promulgating the teachings of Aristotle. The influence of Averroes was wide and lasting, though also disturbing. Renan has said that St. Thomas Aquinas, though combatting his doctrine in many points, yet, as a philosopher, owed almost everything to Averroes.

Averroes opposed the mysticism of Algazel and his followers, and supported the search for truth through human reason. He appears to have had more faith in Aristotle than in revelation, although the two were regarded by him as in no wise opposed. The position and function of imagination in his system is indicated in the following excerpts from Renan (7):

The soul is not divided according to the number of individuals, for it is one in Socrates and in Plato. For intellect has no individuality; individuation comes only from sensibility. . . . The individual soul perceives nothing without imagination. Just as sense is not affected except in the presence of the object, so the soul does not think except in presence of the image.

The conception of reason as active intellect is a cardinal tenet in his teaching. The chief end of man (cf. Renan 95) is to enter into an ever closer union with active intellect. To find and know scientific truth is to enjoy immortal bliss here and now. The man who is absorbed in such activity need not go to heaven; he is there.

Active intellect is one and the same in all men, said Averroes. "*Quod intellectus hominum est unus et idem numero.*" This was one of the prime offenses for which he was condemned in 1269 by the bishop of Paris. One outraged critic asked: "What! is the mind of Judas the same as the mind of St. Peter?" That, of course, sounds terrible. The Christian view was that each human being is created with an individual intelligence, and that this is individually immortal. The "unity of intellect" seemed to negate human personality. These two views are not fatally opposed. Unity is always unity of something, it is not just unity. And variety just as necessarily implies some oneness to give it meaning. There is no harmony without parts and no parts without a whole. The ideas of One and Many are correlatives. If there is such a thing as psychology, a science of mind, it must be based on the assumption that minds are essentially the same. You cannot have a science of particulars.

Ibn Khaldun (1332-1406) (5), born in Tunis of a distinguished family,

was the author of a treatise on Universal History. His translator, Issawi, says of the Prolegomena to this work that it "represents the most comprehensive synthesis of the Human Sciences ever achieved by the Arabs."

Many judicious psychological observations are included in Ibn Khaldun's discussions of social and political life. And he offers valuable advice on education. There are, too, some explicit statements on the various mental operations, including imagination. Following is a description of abstraction in a concrete setting (Issawi p. 162):

Travelling in pursuit of knowledge and for the purpose of meeting new teachers makes learning more perfect. This is because men acquire their knowledge, characteristics, virtues, and opinions either through book learning or through direct contact and precept. Habits and skills acquired through contact and precept are, however, stronger and more deeply rooted; hence the more numerous the teachers with whom the student has been in direct touch, the more deeply rooted is his skill.

Moreover technical terms and methods are confusing to students, who are often inclined to think that these terms or methods form an integral part of the science. The only way to remove this confusion is for the student to have direct experience of different methods, under different teachers. Such contact with learned men and teachers will give him more discrimination, owing to the diversity of terms and methods which he will come across. He will then be in a position to abstract the science from any particular set of terms and to understand that terminologies and methods of study are merely means to be used in the acquisition of science.

This is a good example of that "dissociation by varying concomitants" which modern psychologists recognize as a basic mental achievement.

Probably the clearest contribution of Ibn Khaldun to the study of psychology lies in his expositions of human habit and character as revealed in the course of human history. What man imagines and wills is shown to be conditioned by climate, tribal organization, financial and economic opportunity, and social institutions, and language. This is a type of psychology hardly yet well developed in our own time.

REFERENCES

1. ALGAZEL. *Dogmatica, Moral, Ascetica*: por Miguel Asin Palacios. Zaragoza: Comas Hermanos, 1901.
2. ASIN, M. P. *Abenmassarra y su escuela*. Madrid: 1914.
3. DE BOER, T. J. *Geschichte der Philosophie im Islam*. Stuttgart: Fr. Frommanns Verlag, 1901. Pp. 191.
4. GUILLAUME, A. In *The Legacy of Islam*. Chapter on Philosophy and Theology. Oxford: Clarendon Press, 1931.

5. IBN KHALDUN. *An Arab Philosophy of History: Selections from the Prolegomena of Ibn Khaldun of Tunis*. Translated and arranged by Charles Issawi. London: Murray, 1950. Pp. xiv and 190.
6. O'LEARY, D. L. *Arabic Thought and its Place in History*. London: Kegan Paul, 1922. Pp. 320.
7. RENAN, E. *Averroes et l'averroësisme*. (3rd ed.) Paris: Michel Levy Frères, 1866. Pp. xvi and 486.

516 Woodruff Avenue
Los Angeles 24, California

PERFORMANCE ON A TRACKING TASK AS A FUNCTION OF POSITION, RADIUS, AND LOADING OF CONTROL CRANKS: I. STATIONARY TARGETS*

Department of Psychology, University of Rochester

ARNOLD A. GERALL, PHILIP B. SAMPSON, AND S. D. S. SPRAGG¹

A. INTRODUCTION

A task of psychologists working in the area of human engineering is to determine functional relationships between an operator's performance and the physical characteristics of the machine parts which he manipulates. These relationships or laws when found will enable the engineer to design equipment so that the operator can perform the necessary tasks with a greater degree of efficiency and precision. They will also increase the psychologist's understanding of the capacity and efficiency of the human organism in various perceptual-motor situations. The number of laws available at present is very small and a great deal of basic research is still necessary. The general purpose of the experiments to be reported in this series of articles was to obtain information concerning the relationship between operators' performance and several physical variables associated with the controls of a tracking task. The results might aid in the design of equipment for improving human performance and in the description of man's capacity to respond under various situational demands.

B. STATEMENT OF THE PROBLEM

The purpose of the present study was to investigate the relationships between performance on a simple two-hand target acquisition task and the following three variables associated with the controls: (a) their position, (b) their radius, and (c) the force necessary to rotate them.

C. SUBJECTS

One hundred eighty recent military enlistees served as subjects. Each was assigned randomly to one of 18 experimental conditions. Anthropometric data were secured for the last 90 subjects. Analyses of the distri-

*Received in the Editorial Office on September 23, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹This research was carried out on Contract No. DA-49-007-MD-326 between the University of Rochester and the Medical Research and Development Division, Office of the Surgeon General, Department of the Army.

butions of heights and weights indicated that the means of each of these measures were approximately the same for all the experimental conditions. The mean height of the subjects was 69.15 inches with a standard deviation of 2.3. The mean weight was 152.92 pounds with a standard deviation of 29.4. These values are similar to those for a military population reported in the *Handbook of Human Engineering Data for Design Engineers* (8).

D. APPARATUS

A detailed description of the apparatus has been presented in the earlier report (1). In brief, it consisted of display, control, and recording units. The display unit was a dual beam oscillograph. One beam of the oscillograph was focused to a small spot and was used as the target. The other beam was converted into a circle and was used as the target follower. By means of an automatic programming device a stationary target was projected on the cathode ray tube (C.R.T.) in various locations. The operator controlled the movement of the target follower by rotating two adjustable bar cranks. The left crank governed the elevation of the circle and the right crank the azimuth of the circle. Each crank was mounted on the shaft of an electromagnetic clutch. By varying the current through the coils of the clutches, the coulomb friction supplied by these units provided torque loads from 4 inch-pounds to 90 inch-pounds.

Continuous records of the operator's performance were obtained on a Brush Oscillograph. The circuits of this unit were designed so that the movement of each control was recorded separately. Each record could be analyzed as to: (a) the time elapsed before the operator made an initial movement of the controls after the target appeared; (b) the duration and direction of movement of the controls during the approach to the target; and (c) any adjustive movements that occurred after the target was reached.

The subject sat in a seat which was adjusted in height so that he could comfortably place his forehead against a guard. This head guard was 12 inches from the surface of the C.R.T. At this distance, one revolution of the cranks produced a movement of the target follower of one degree of visual angle. Twenty revolutions of the cranks moved the target follower four inches on the C.R.T. A photograph of the apparatus that shows the position of the operator's chair, the cranks, and the display is provided in Figure 1.

E. PROCEDURE

At the beginning of each trial, the target and target follower would appear displaced from each other on the face of the C.R.T. The task of the sub-

ject was to rotate the controls until the target follower encircled the target. The target appeared on the C.R.T. in eight positions. Each position or displacement was displayed for 30 seconds. Then the circle and spot disappeared and 15 seconds elapsed before the next presentation of the target.

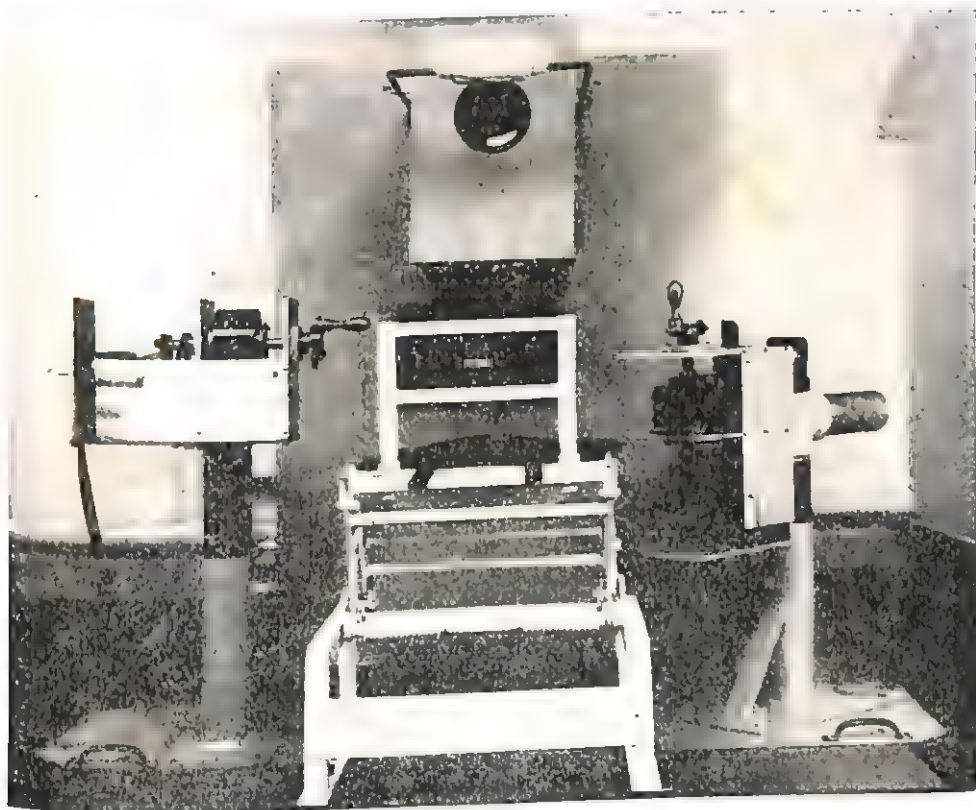


FIGURE 1
PHOTOGRAPH OF THE TRACKING APPARATUS SHOWING THE OPERATOR'S LOCATION IN
RELATION TO THE CONTROLS AND THE DISPLAY: THE CRANKS ARE
PLACED IN POSITION 4

After the operator had circled the target at all eight positions, he was given a rest of approximately one minute. At the end of this period the same eight displacements were presented to him. Each operator was given four such practice series. Hence, the extent of learning or practice, as represented by the scores obtained during the succeeding series, could also be analyzed.

A factorial design which incorporated all combinations of the experimental variables was used in this study. The sizes of the crank radii were 2, 4, and 6 inches and the magnitudes of forces necessary to rotate the cranks were

2, 8, and 14 pounds. The positions of the controls were designated as *A* and *B*. In Position *A*, the left crank was rotated in a vertical plane parallel to the sagittal plane of the body and controlled the movement of the target follower in the vertical direction on the scope. The right control was rotated in a horizontal plane. This control unit was located at a 45° angle to the frontal plane of the body and governed the horizontal movement of the target follower. The only difference between the two positions was the location of the right crank. For Position *B*, it was rotated in the vertical plane 45° to the sagittal plane. The first 90 subjects were tested on Position *A* and the second 90 subjects on Position *B*.

F. RESULTS

1. *Total Time to Reach the Target Scores*

The total time required by the operators to reach the target was recorded for every trial during the experimental period. In the analyses to be discussed below, an average time was computed for each subject from the scores of the eight trials in each of the four series. Accordingly, four scores were available for each operator which represented his proficiency at each stage of practice. The first analysis performed on these scores was a test of the significance of the obtained difference in performance that might be attributed to the two positions of the controls. This analysis showed that the operators who used Position *A* reached a significantly higher level of proficiency than those who used Position *B*. (The 5 per cent confidence level was used for this and all subsequent tests.) In Figure 2, the groups' performance on the two positions is plotted for each force used in the experiment. It should be noted in this graph and those to follow that lower total time to reach the target score is indicative of a higher level of performance on this task.

A more detailed analysis of the effect of the two positions upon performance revealed an unexpected and complicated result. Summary tables of the analyses of variance completed for each position separately are given in Tables 1 and 2. It is apparent from a comparison of these results that the difference in position is associated with differences in interaction effects. For Position *A*, there were no interaction effects between series (i.e., practice) and any of the other variables. However, for Position *B*, there was a reliable interaction between series and the crank variable. Figures 3 and 4 illustrate the improvement in performance with practice of each crank on the two positions. The implication of this finding is that for Position *B* a statement concerning the relative effectiveness of different sizes of cranks

would also have to include a qualification concerning the degree of training of the operators who used them in this position. In other words, the interaction between cranks and series restricts the generalizability of the results obtained with Position *B*. Because of the limited practical application of

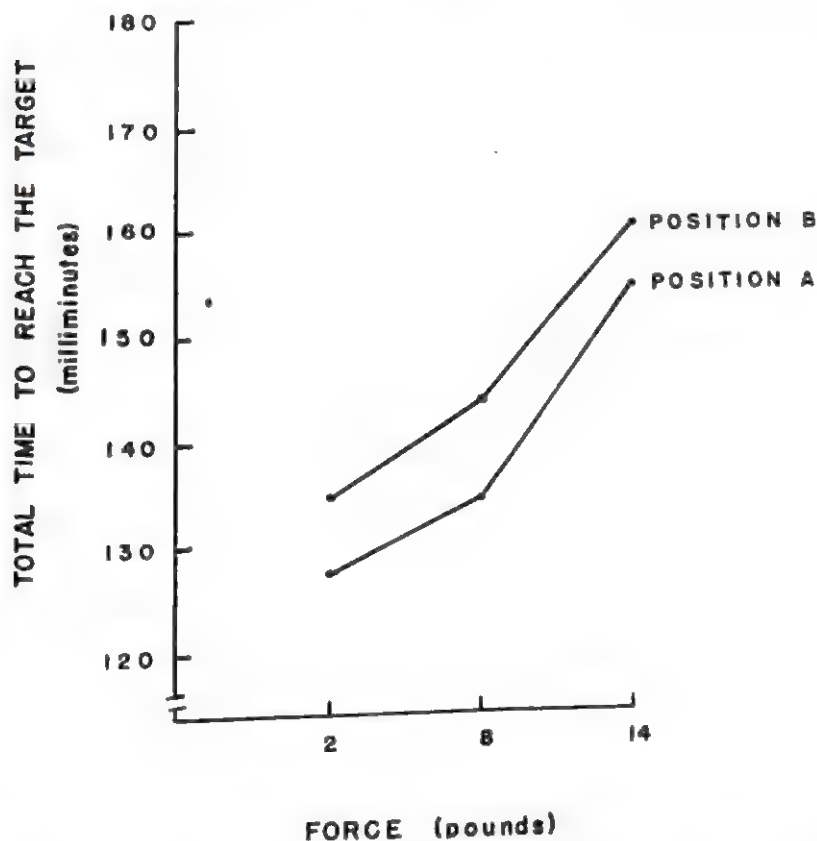


FIGURE 2
PERFORMANCE OF TWO GROUPS OF OPERATORS ON DIFFERENT CONTROL POSITIONS AS A
FUNCTION OF THE FORCE NECESSARY TO TURN THE CONTROLS
($N = 90$ for Each Group)

the results secured thus far on Position *B* and because this position was found to be significantly inferior to Position *A*, attention in this report will be focused mainly upon the data from Position *A*.

According to the results shown in Table 1 the crank by force interaction was not reliable for Position *A*. In Figure 5, the relation between the three magnitudes of forces and performance is depicted for each crank. If there were no interaction between crank and force all the curves would be approxi-

mately parallel to each other. Obviously, the curves do not maintain a parallel course. It is probable a real interaction exists between the two factors and that it would have been found to be statistically significant if a larger number of subjects had been used in this experiment. This interaction

TABLE 1
ANALYSIS OF VARIANCE FOR TOTAL TIME TO REACH TARGET WITH CONTROLS IN POSITION *A*

Source	<i>df</i>	<i>ms</i>	<i>F</i>
Crank	2	1,776	0.5
Force	2	23,283	6.9**
<i>FxC</i>	4	5,491	1.6
Error (b)	81	3,381	
Series	3	48,539	233.2**
<i>SxF</i>	6	346	1.7
<i>SxC</i>	6	294	1.4
<i>SxCxF</i>	12	87	0.3
Error (w)	243	208	

*Beyond the 5 per cent level of confidence.

**Beyond the 1 per cent level of confidence.

TABLE 2
ANALYSIS OF VARIANCE FOR TOTAL TIME TO REACH TARGET WITH THE CONTROLS
IN POSITION *B*

Source	<i>df</i>	<i>ms</i>	<i>F</i>
Crank	2	4,790	1.5
Force	2	20,304	6.5**
<i>CxF</i>	4	8,203	2.6 ^a
Error (b)	81	3,140	
Series	3	42,116	205.4**
<i>SxF</i>	6	176	0.8
<i>SxC</i>	6	959	4.7**
<i>SxCxF</i>	12	220	1.1
Error (w)	243	205	

*Beyond the 5 per cent level of confidence.

**Beyond the 1 per cent level of confidence.

signifies that the effect that cranks with a given radius have upon performance depends, at least in part, upon the force which the operator has to exert to turn the cranks.

The preceding analyses and graphs consistently show that if there were any change in performance with an increase in coulomb friction present in the control system and a consequent increase in force requirement for the operator, it always was manifested as a decremental effect with Position *A*. However, the rate of the change of the decrement was not the same for the three radii of cranks. For a given increase of force, the associated decrement in tracking speed was greatest for the 2-inch crank and least for the

6-inch crank. This result is reflected in the slopes of the curves presented in Figures 5 and 6.

Figure 6 also reveals that each crank radius has associated with it a par-

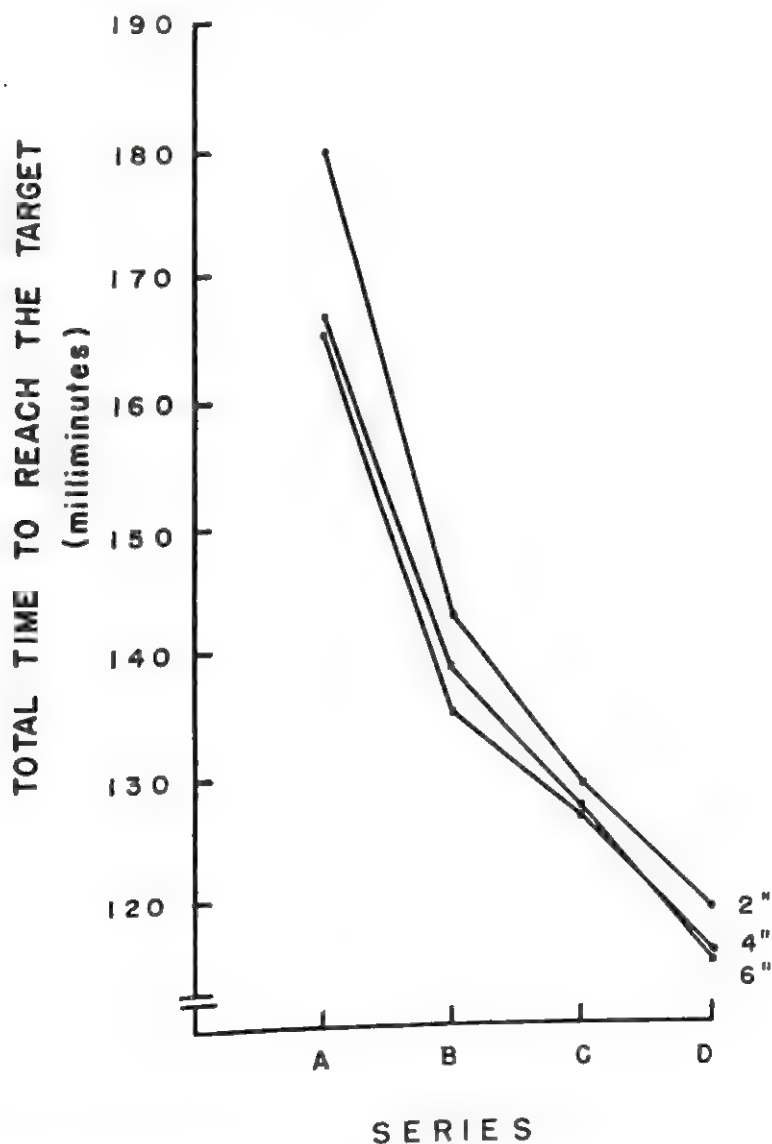


FIGURE 3
CHANGE IN THE PERFORMANCE OF OPERATORS WHO USED DIFFERENT CRANK RADII
WHEN THE CONTROLS WERE PLACED IN POSITION A DURING
THE FOUR PRACTICE SERIES

ticular range of torques over which it leads to more proficient performance than another crank radius. Furthermore, the extent of the range appears to be related to the size of the crank.

The variability of performance associated with each experimental condition

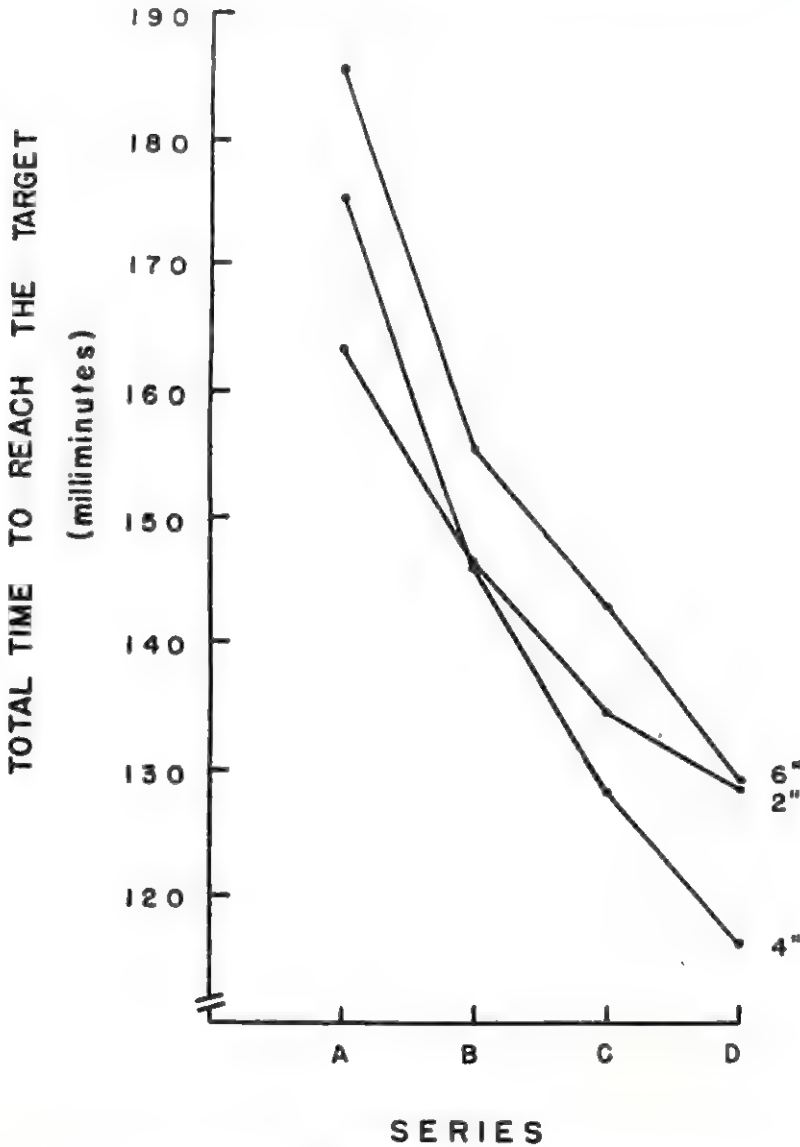


FIGURE 4
CHANGE IN THE PERFORMANCE OF OPERATORS WHO USED DIFFERENT CRANK RADII
WHEN CONTROLS WERE PLACED IN POSITION B DURING
THE FOUR PRACTICE SERIES

was also analyzed. The largest variability among the operators was found for the 14-pound, 2-inch condition. A comparison of the over-all difference in variability of performance between the groups, however, was found not to be reliable.

For the reasons stated previously, emphasis was placed on the results obtained on Position *A*. It would be informative at this point to consider briefly some of the data obtained with Position *B*. The form of the func-

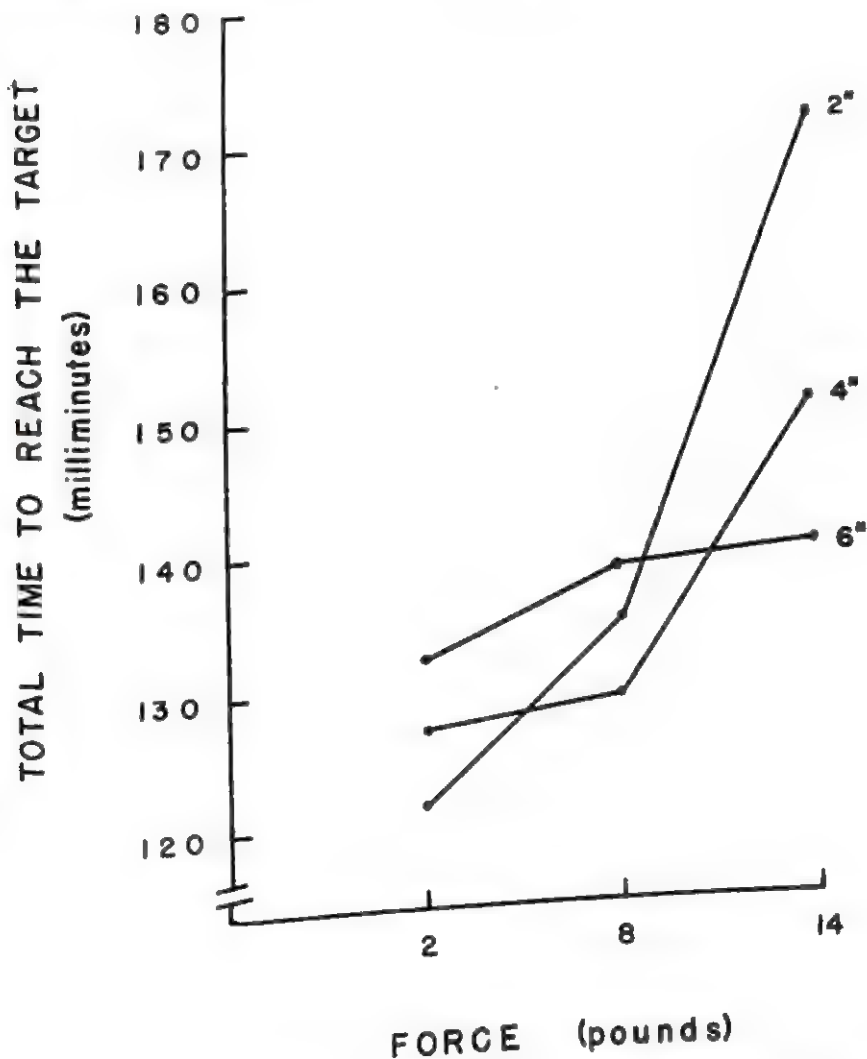


FIGURE 5
PERFORMANCE OF OPERATORS WHO USED DIFFERENT SIZES OF CRANK RADII AS A FUNCTION OF THE FORCE NECESSARY TO TURN THE CRANKS WHEN PLACED IN POSITION *A*

tion for the 2-inch crank was essentially identical for the two positions. It is probable that the conclusions drawn for the 2-inch crank will apply equally to either position. Another similarity between the two conditions was shown for the 6-inch crank when it was coupled with 2- and 8-pound requirements. However, when 14 pounds were required to rotate the 6-inch crank in Posi-

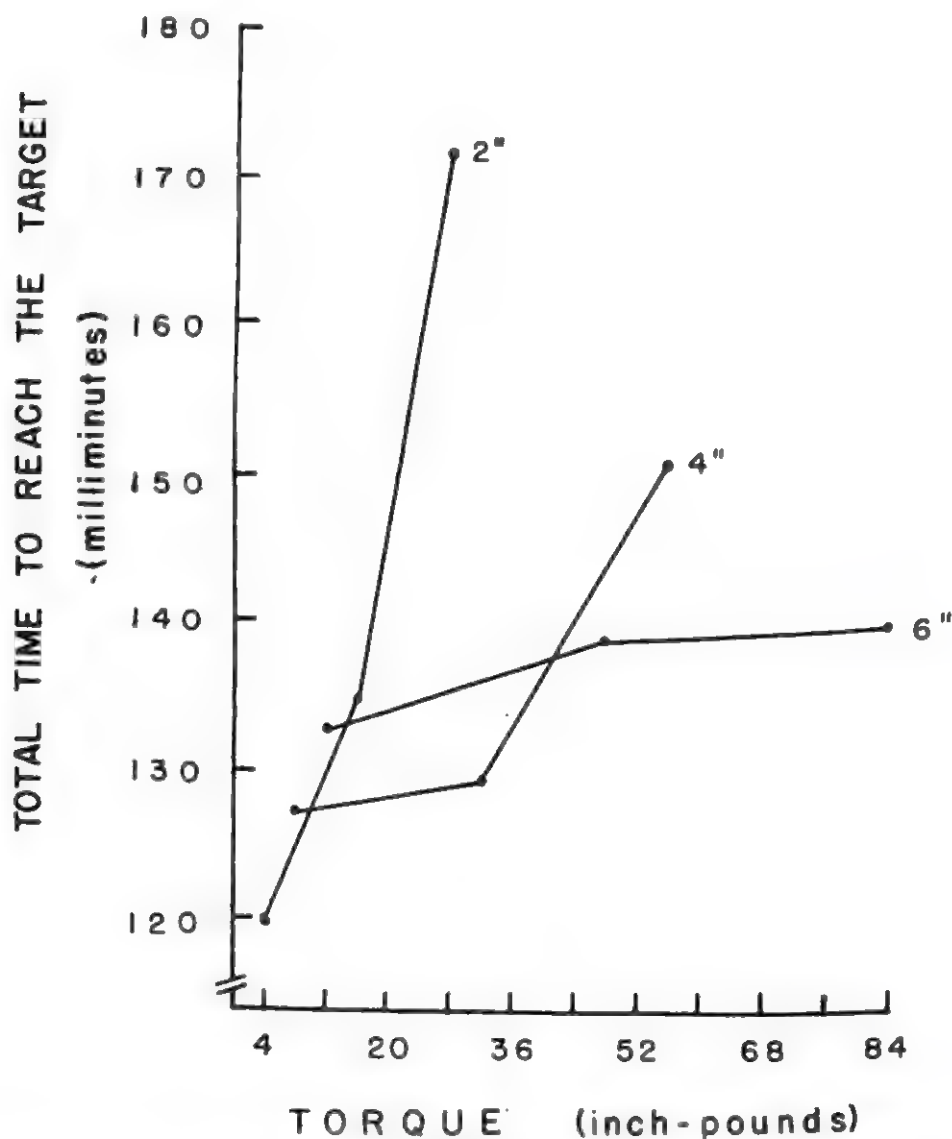


FIGURE 6
PERFORMANCE OF OPERATORS WHO USED DIFFERENT SIZES OF CRANK RADII AS A FUNCTION OF THE TORQUE DEMAND IMPOSED UPON THE CRANKS WHEN PLACED IN POSITION A

tion *B*, it resulted in longer acquisition times than those associated with the same condition with Position *A*. The most outstanding discrepancy between the two positions was associated with the 4-inch crank. With this crank in Position *B*, the group that had the 14-pound condition obtained lower tracking scores than the groups that had the 2- and 8-pound conditions. The comparable group on Position *A* had the longest time to reach the target score in relation to the other two groups. Also, as shown in Figures 3 and 4, the rate of change in performance with practice for the 4-inch crank relative to the other groups was greater with Position *B* than with Position *A*. Finally, the over-all variability of Position *B* was greater than that of Position *A*.

2. Component Responses

In addition to total time to reach the target scores, starting, travel, and adjustment times, and error scores were secured from the graphic records of the subjects' performance. Starting time was calculated from the interval between the presentation of the target and target follower and the first measurable indication of movement on the record. As such, it is a measure which contains some movement time of unknown duration. Also, the starting time scores used in the analysis to be discussed below were obtained from the hand which yielded the longest time for each target acquisition. An analysis of variance of these scores is presented in Table 3. According to

TABLE 3
ANALYSIS OF VARIANCE OF STARTING TIME

Source	<i>df</i>	<i>ms</i>	<i>F</i>
<i>Between</i>			
Crank	2	1,108	3.1*
Force	2	2,986	8.5**
Position	1	2,044	5.8*
<i>CxF</i>	4	1,241	3.5**
<i>CxP</i>	2	176	
<i>FxP</i>	2	212	
<i>CxFxP</i>	4	231	
Error	162	352	
<i>Within</i>			
Series	3	5,117	76.4**
<i>CxS</i>	6	30	
<i>FxS</i>	6	147	2.2*
<i>PxS</i>	3	31	
<i>SxCxF</i>	12	46	
<i>SxPxP</i>	6	6	
<i>SxPxS</i>	6	31	
<i>SxCxFxP</i>	12	52	
Error	486	67	

*Beyond the 5 per cent level of confidence.

**Beyond the 1 per cent level of confidence.

this analysis, the two positions used in this experiment had a differential effect upon starting time scores. In Figure 7a, the differences between Positions *A* and *B* are shown as a function of the three crank radii. Operators who used Position *A* had a longer starting time than those who used Position *B*. It should be noted at this point that total time to reach the target scores were longer for Position *B* than they were for Position *A*. Thus, the position which afforded the longer time to reach the target, nevertheless, enabled the operators to manifest the shorter starting time.

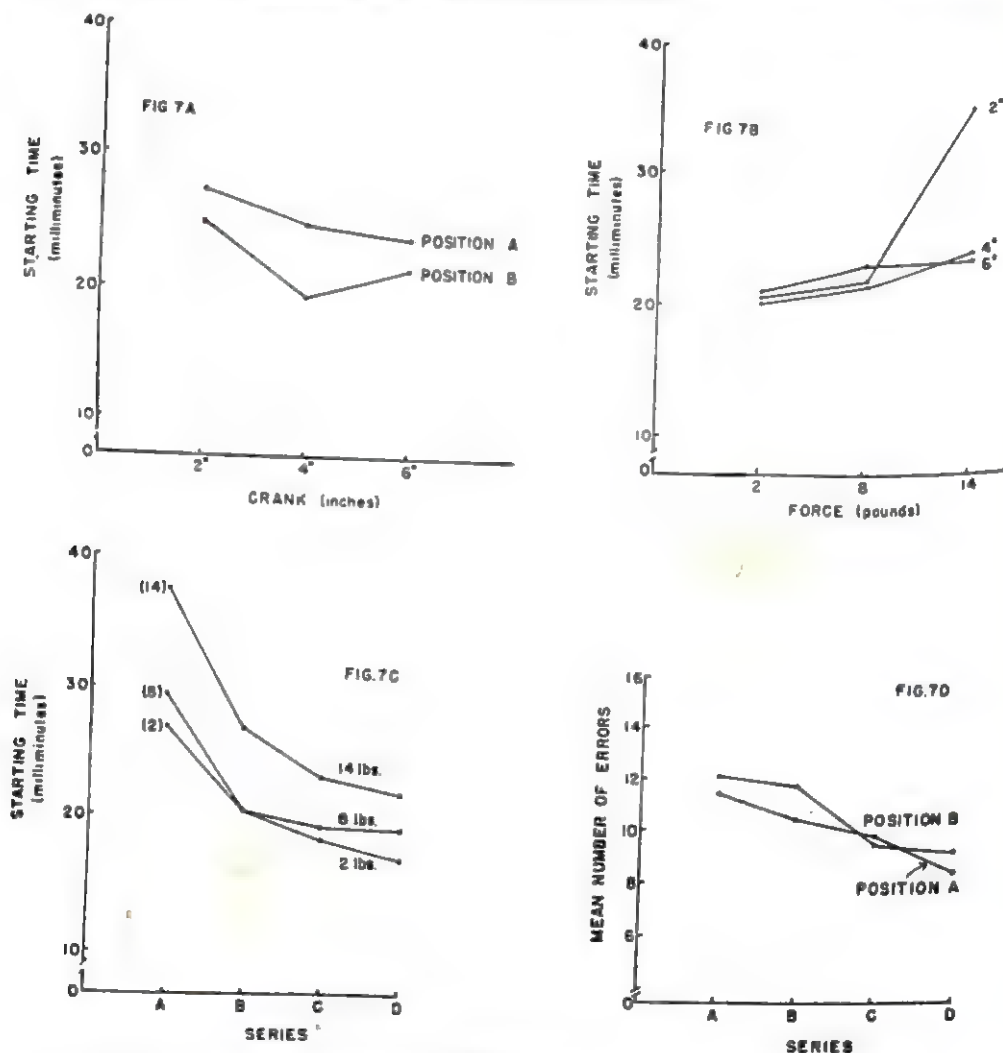


FIGURE 7

- Starting time as a function of crank radius, Positions *A* and *B*.
- Starting time as a function of force, for each crank radius.
- Starting time as a function of sequence, for each force.
- Mean number of errors as a function of sequence, Positions *A* and *B*.

The relationship between starting times and forces for each of the three crank radii is shown in Figure 7b. Starting times increase as a function of force but the rate of increase was not the same for each crank radius. This is illustrated by the curves in Figure 7b and the significant interaction obtained between crank and force variables in the analysis of variance. With an increase in force from 2 to 14 pounds, the increase in starting times was greatest for the 2-inch crank and least for the 6-inch crank.

The analysis of variance also indicated that starting time decreased as a function of practice. This result is presented in Figure 7c. There was a slight tendency for the curve representing starting time when eight pounds of force were required to have a different slope than the curves for the 2 and 14 pounds. The force by series interaction, which is a quantitative estimate of the similarity of the rate of change of these functions, was found to be statistically reliable. The degree of deviation from parallel courses of the curves, although statistically reliable, is small and probably does not have immediate practical significance.

Travel time was measured as the interval between the end of starting time and the point at which the target was first reached. It is comprised mainly of movement responses of the operators. The travel time component was the largest element of total time. As might be expected, travel time results are essentially similar to those found for total time to reach the target.

The data for adjustment times were of interest primarily in that their frequency of occurrence was extremely small. No systematic relationships between adjustment time and position, radius, or force requirements of the cranks were found.

Errors were defined in this study as movements made in a direction away from the target. An analysis of variance for the total number of errors made with either control indicated that only series or practice seemed to affect reliably the number of errors made. Fewer errors were made with increase of practice on this task. There is a trend, although it is not statistically significant at a level established previous to obtaining the results, for fewer errors to be made with Position *A* than with Position *B*. But it is of interest that the difference was as small as it was. A graph showing the decrease in errors with practice and the difference in the number of errors recorded for each position is shown in Figure 7d.

G. DISCUSSION AND SUMMARY

The purpose of the present study was to investigate the relationship between performance of operators on a simple two-hand target acquisition task and several aspects of the control system associated with the tracking device.

The variables studied were the position, radius, and force requirement of the controls. There were two positions, three sizes of crank radii, and three magnitudes of force requirements. All combinations of the selected dimensions of these variables were investigated and analyzed according to procedures appropriate to factorial designs of experiments. The primary measure of the operator's performance on this task was total time to reach the target, and the discussion will be based upon this measure unless specified otherwise.

An analysis of the influence of the different positions of the controls revealed that this variable had several significant effects upon performance. Position *A* enabled the subjects to obtain a higher and less variable level of performance than did Position *B*. A single cause of the superiority of Position *A* cannot be identified at this time. It is possible that the location of the cranks in Position *A* enable the operators to use their arms with greater speed and force than does Position *B*. However, the difference may not be due to physiological factors of the motor system but to perceptual-motor influences. Several experiments have been reported recently which evaluated many display-control relationships for two-hand tracking tasks (2, 6). The results of these experiments suggest that the arrangement of the controls designated as Position *A* in this study permits the operator to reach a higher degree of proficiency than would controls placed in an arrangement such as Position *B*, because the relationship between the plane and direction of rotation of the controls of Position *A* are more "natural" and "continuous" with the movement of the target and target follower on the display than are those of Position *B*. The tendency for fewer errors to be made with Position *A* supports the hypothesis that Position *A* is the better of the two positions because of its control-display arrangement. If the perceptual-motor relationship is the more important factor determining the difference between the two positions, then with tasks that do not require tracking, such as a pure rotary task, Position *A* might not be more effective than Position *B*.

At first glance, the starting time data do not appear consistent with the suggestions made above. It would be expected that the position with the better control-display relationship would lead to the shorter starting times. However, operators using Position *A*, which had the more natural control-display relationship, had the longer starting times. This result might be due to the fact that starting times as measured in this experiment had a large motor or movement component contributing to its duration. If Position *B* were the better arrangement for the application of force, then the starting times might be shorter when this position was used because of this advantage.

The influence of practice upon the performance of the operators was also different for the two positions. When Position *B* was used, a triple interaction between the crank, force, and practice variables were found. This interaction indicates that the relative effects of the crank and force variables shifted during the practice series. At present, no adequate explanation can be offered by us to account for the nature or cause of this interaction. No interaction between practice and the crank and force variables were found when Position *A* was used. Thus, the relative effects of the crank and force variables appears to be more independent of the degree of experience of the operators with this control arrangement.

With Position *A*, the influence of the different crank radii upon performance was found to depend upon the force necessary to turn them. The same result can be expressed in terms of the torque load of the control system as well as the force demand upon the operator. For low values of torque, the 2-inch crank was more effective than the others. The usefulness of a crank with a small radius, however, decreases rapidly as the torque of the system increases. Therefore, cranks with radii of two inches or less should be used only when the torque demand of the system is not large (probably not over 10 inch-pounds). The range of optimal usefulness of the 4-inch crank in this study was approximately between 10 and 40 inch-pounds. After 40 inch-pounds, the 6-inch crank seems to be superior to cranks with smaller radii.

Although a relatively large number of subjects was used in the entire experiment, it is premature at this time to specify the exact range of torques over which a given crank is more effective for the general population. A larger sample of subjects would be necessary to achieve this desirable goal. In lieu of such a normative study, these data are useful in suggesting that such ranges of optimal performance for each crank radius do exist and in providing very approximate values of the boundaries of the ranges. Also, it is important to note that in systems which have variable torque demands that cranks with larger radii will lead to more nearly constant performance over a wide range of force values than cranks with smaller radii.

The friction imposed upon the control of the tracking device employed in this study was essentially independent of the velocity of rotation. This type of friction is usually designated as coulomb friction. According to the data obtained with Position *A*, an increase in coulomb friction causes a decrease in performance for all sizes of cranks. Helson (3, 4) has studied the influence of coulomb friction upon performance with uni-dimensional compensatory tracking tasks both using one and two hands. In almost all in-

stances, Helson reported that larger magnitudes of friction increase both tracking errors and irregularity of performance. The results of this experiment complement those obtained by Helson. Hick (5) compared the effects of friction upon one hand tracking performance under conditions of jolting and non-jolting of the control. He reported that friction was favorable in jolting conditions for tracking involving short to-and-fro corrective movements but unfavorable for the same task under non-jolting conditions. Reed (7) studied the effect of friction upon rate of rotating cranks of various radii. Friction always acted to decrease the speed of turning for a given crank radius. Thus, the data available at present seem to be consistent with the conclusion that coulomb friction is detrimental to the efficient use of cranks when the task involves rotation alone, tracking under relatively stable conditions of the controls, or two-hand acquisition of stationary targets.

REFERENCES

1. GERALL, A. A., SAMPSON, P. B., & SPRAGG, S. D. S. An electronic apparatus for studying tracking performance. *Amer. J. Psychol.*, 1955, **68**, 297-305.
2. GREEN, R. F., ANDREAS, B. G., NORRIS, EUGENIA B., & SPRAGG, S. D. S. Performance on a following tracking task (the SAM Two-Hand Coördination Test) as a function of the continuity of the plane and direction of movement of the control cranks and target follower. *J. of Psychol.*, 1955, **40**, 403-410.
3. HELSON, H. Inertia, friction, and diameter in handwheel tracking. National Defense Research Committee, Division 7, Report to the Services No. 74, Foxboro Company, 1943.
4. ———. Design of equipment and optimal human operation. *Amer. J. Psychol.*, 1949, **62**, 473-497.
5. HICK, W. E. Friction in manual controls with special reference to its effect on accuracy of corrective movements in conditions simulating jolting. Medical Research Council, Unit of Applied Psychology, The University, Cambridge, Technical Report MRC 47/386, APU 18, June 1945.
6. NORRIS, E. B., & SPRAGG, S. D. S. Performance on a following tracking task (modified SAM Two-Hand Coördination Test) as a function of the relations between direction of rotation of controls and direction of movement of display. *J. of Psychol.*, 1953, **35**, 119-129.
7. REED, J. D. Factors influencing rotary performance. *J. of Psychol.*, 1949, **28**, 65-92.
8. Tufts College Institute for Applied Experimental Psychology, *Handbook of Human Engineering Data for Design Engineers*. Office of Naval Research, Special Devices Center, Port Washington, L. I., N. Y., Technical Report SDC 199-1-1, 1949.

Department of Psychology
University of Rochester
Rochester 3, New York

PERFORMANCE ON A TRACKING TASK AS A FUNCTION OF POSITION, RADIUS, AND LOADING OF CONTROL CRANKS: II. MOVING TARGETS*

Department of Psychology, University of Rochester

ARNOLD A. GERALD, PHILIP B. SAMPSON, RUSSEL F. GREEN, AND
S. D. S. SPRAGG¹

A. INTRODUCTION

In a previous study, variations in the positions, radius, and loading of the controls were found to influence the performance of operators of a target acquisition task. Since there are many types of tracking devices commonly used in laboratory investigations and field situations, additional analyses and research appeared to be necessary before the generality of the results could be ascertained.

At the beginning of each trial of the target acquisition task employed in the aforementioned study, the target and target follower appear in separate locations on the face of a cathode ray tube. Each operator is instructed to rotate the controls so that the target follower reaches and encircles the stationary target in the shortest possible time. The measure of the operator's proficiency, therefore, is the time required to reach the target. The level of proficiency achieved in this target acquisition task depends upon, in large part, the ability of the operator to coördinate and maximize the speed of rotation of the controls. They must coördinate the speeds of rotation of the controls if the target follower is to be moved along the shortest path to the target. For example, for one target displacement, the "x" crank may have to be rotated twice as fast as the "y" crank if the target follower is to take a straight line course to the target. For other target displacements, the ratio of rotating the cranks changes according to the position of the target in the "x" and "y" directions relative to the target follower. The absolute rate of rotating each crank determines the speed at which the target follower moves in the intended direction. Observations of the behavior of operators on the target acquisition task under discussion seem to indicate that they

*Received in the Editorial Office on September 23, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹This research was carried out on Contract No. DA-49-007-MD-326 between the University of Rochester and the Medical Research and Development Division, Office of Surgeon General, Department of the Army.

were more intent upon maintaining speed rather than coördination of the controls.

In another common type of tracking, called following tracking, the target is in motion during most of the trial. Operators engaged in following tracking are required to keep the target follower over a moving target as long as possible. Accordingly, the primary measure of performance is total time on target. Mastery of the task occurs when the operator is able to coördinate the action of both hands in response to the perceived movements of the elements on the display. Absolute speed of rotating the controls is largely determined by the speed of movement of the target. In most following tracking tasks, the target speed is such that maximum effort or speed of rotation of the controls is not required of the subjects. Thus, ability to coördinate the movement of their two hands appears to determine performance in following tracking while ability to maintain a high speed of rotation of the controls seems to regulate performance in target acquisition tasks. If this analysis is correct, then these two tasks depend, at least in part, upon different skills. The experiment to be reported here, was intended to discern whether the relationships found between performance on a target acquisition task and three physical control variables would be relevant to a following tracking task.

B. PURPOSE

It was the purpose of this experiment to determine the influence of two different positions, three sizes of radii, and three magnitudes of loadings of control cranks upon performance of operators of a following tracking task.

C. SUBJECTS

One hundred fifty-three military trainees served as subjects in this experiment. The first 108 subjects performed with controls in Position *A*. The remainder of the subjects served with the controls placed in Position *B*. Each subject was assigned randomly to one of nine radius and force conditions. There were no statistical differences among the average weight or height of the groups assigned to the 18 experimental groups.

D. APPARATUS

The programming unit of the tracking device described in detail in a previous report (1) was modified for the purposes of the present study. Originally, the programming unit controlled presentation of the stationary targets and the temporal intervals between trials. At the onset of each trial in the target acquisition situation, the target would appear displaced from the

target follower. The extent of the displacement was determined by selecting two voltages, one for the "y" component, and the other for the "x" component, from two banks of fixed resistors in series. In order to convert the programmer so that a moving target could be presented, two motor driven, high resistance potentiometers were switched in parallel with the resistances that were to be selected for a given displacement. The target, therefore, moved between each successive position formerly used in the target acquisition task. This also means that approximately the same number of revolutions of each control was required in both situations. The rate of movement of the target was $1^{\circ}41'$ of visual angle per second or $1/3$ of an inch per second on the face of the cathode ray tube.

The programmer was also modified to shorten the duration of a trial and the inter-trial interval. In the previous study each trial lasted for 30 seconds and 15 seconds elapsed between trials. For the present study, each trial continued for 8.3 seconds and the inter-trial interval was 21.7 seconds.

A Standard Electric Clock was used to record time on target. Whenever any part of the target follower touched the target, the clock was activated by an electronic relay. The same electronic relay was used in the previous study but its contacts were re-wired for recording time on target. All the other aspects of the equipment were the same as described in earlier reports (1, 2) and the reader is requested to refer to them for additional information.

E. PROCEDURE AND DESIGN

On the first trial of every series, the target and target follower appeared in the same location on the face of the cathode ray tube. Almost immediately after it appeared, the target began to move at an approximately constant velocity and in a straight line for a predetermined distance. The task of the operator was to keep the target follower, which was a circle, over the target spot as it moved. When the target reached the end of its course, it remained stationary for several milliminutes before it disappeared from the face of the scope. After the 21.7 seconds of the inter-trial interval elapsed, both the target follower and the target re-appeared where the operator left them at the end of the previous trial. The target then started to move at the same velocity as before but in a new direction. All of the eight courses differed from each other in the ratio of vertical and horizontal distance traversed during a trial.

After the operator tracked all eight courses in a series, he was given a one-minute rest. At the end of this period, the same eight target courses were repeated in the same order. Each operator was given four series and then

permitted to rest for approximately 25 minutes before he was re-tested on a second set of four series.

The design and experimental conditions used in the previous study were reproduced in this experiment. A factorial design which incorporated all combinations of the experimental variables was used. The sizes of the crank radii were 2, 4, and 6 inches and the magnitudes of force necessary to rotate the cranks were 2, 8, and 14 pounds. The positions of the controls were designated as *A* and *B*. In Position *A*, the left crank was rotated in a vertical plane parallel to the sagittal plane of the body and controlled the movement of the target follower in the vertical direction on the scope. The right control was rotated in a horizontal plane. This control unit was located at a 45° angle to the frontal plane of the body and governed the horizontal movement of the target follower. The only difference between the two positions was the location of the right crank. For Position *B*, it was rotated in the vertical plane 45° to the sagittal plane of the subject. The first 108 subjects were tested on Position *A* and the second 45 subjects on Position *B*.

F. RESULTS

An average time on target score was computed from the eight trials of each practice series for each subject. The basic 18 experimental conditions, therefore, were replicated eight times in this design. Of course, the replications represent practice effects or repeated measurements on the same subject. The results of the analysis of variance performed on these data are presented in Table 1. The only statistically significant effect was associated with the series or practice variable. There was a tendency for the operators who used Position *A* to obtain higher time on target scores than those who used Position *B*. Also, there was a trend for performance level to decrease with an increase in force required to rotate the controls. Neither of these tendencies, however, could be held on statistical grounds to be due to anything other than chance variation.

G. DISCUSSION AND CONCLUSION

An effort was made in this study to reproduce the experimental conditions of a previous experiment with one major exception. The target acquisition task used in the previous experiment was converted into a following tracking task. It was the purpose of this experiment to determine whether the variables of position, radius, and loading of the control cranks affect performance of operators of target acquisition and following tracking tasks in a similar manner.

Performance of operators of a following tracking task was found not to be systematically related to the variables of position, radius, and loading of the control cranks. These same variables had been found to influence performance of operators in a target acquisition situation. There was some indication that at least one of the combinations of radius and force demands, the 2-inch-14-pound control, led to a greater decrement in performance than the other controls. These effects, however, were not statistically evident in an over-all F-test and shall have to be demonstrated in future research.

TABLE 1
ANALYSIS OF VARIANCE OF TOTAL TIME ON TARGET

Source	<i>df</i>	<i>ms</i>	<i>F</i>
Crank	2	38.88	
Force	2	3,058.93	
Position	1	624.27	
<i>CxF</i>	4	423.18	
<i>CxP</i>	2	814.39	
<i>FxP</i>	2	778.58	
<i>CxFxP</i>	4	582.19	
Error (b)	135	2,185.21	
		22,191.67	170.40**
Series	7	225.89	
<i>CxS</i>	14	147.22	
<i>FxS</i>	14	55.16	
<i>PxS</i>	7	99.47	
<i>CxFxS</i>	28	130.99	
<i>CxSxP</i>	14	49.02	
<i>FxSxP</i>	14	196.80	
<i>CxFxSxP</i>	28	130.23	
Error (w)	45		

**One per cent level of confidence.

The absence of a statistically significant difference in the performance of the operators of the two positions was not expected. Several studies (3, 6) have shown that the position of the controls in relation to the display influences performance of operators of two-hand tracking tasks. Position *A* had the more natural or compatible correspondence with the display than Position *B*, in this study. It was expected that Position *A* would enable the operators to obtain higher time on target scores than those who used Position *B*. It is possible that no difference was found because the experiment was not sufficiently precise to uncover it. In the present experiment, inter-subject variability was rather large in spite of efforts to keep motivation high and the task simple. If a larger sample of subjects were used or if the inter-subject variability were decreased by either selection or experi-

mental methods, then position effects, it is suggested, would be found to be different to a statistically significant extent.

There is no direct evidence, to the authors' knowledge, which demonstrates that performance on a two-hand tracking task is related to the radius and loading of the controls. However, on the basis of studies dealing with these variables in other types of situations such as: target acquisition (2), single control rotation (7), and unidimensional compensatory tasks (4, 5), it was anticipated that some effect would be found. Perhaps, in the present experiment, the operators were not required to exert themselves sufficiently and thus were able to compensate for any deleterious effects due to certain crank radius and loading conditions. If either the rate of target movement or the duration of the trial were increased, then the variables of crank radius and loading might be important. However, until further work is completed, it is necessary to conclude that the results found for a target acquisition task reported previously do not apply to a following tracking task.

REFERENCES

1. GERALL, A. A., SAMPSON, P. B., & SPRAGG, S. D. S. An electronic apparatus for studying tracking performance. *Amer. J. Psychol.*, 1955, **68**, 297-305.
2. ———. Performance on a tracking task as a function of position, radius, and loading of control cranks: I. Stationary targets. *J. of Psychol.*, 1956, **41**, 135-150.
3. GREEN, R. F., ANDREAS, B. G., NORRIS, EUGENIA B., & SPRAGG, S. D. S. Performance on a following tracking task (the SAM Two-Hand Coördination Test) as a function of the continuity of the plane and direction of movement of the control cranks and target follower. *J. of Psychol.*, 1955, **40**, 403-410.
4. HELSON, H. Inertia, friction, and diameter in handwheel tracking. National Defense Research Committee, Division 7, Report to the Services No. 74, Foxboro Company, 1943.
5. ———. Design of equipment and optimal human operation. *Amer. J. Psychol.*, 1949, **62**, 473-497.
6. NORRIS, E. B., & SPRAGG, S. D. S. Performance on a following tracking task (modified SAM Two-Hand Coördination Test) as a function of the relations between direction of rotation of controls and direction of movement of display. *J. of Psychol.*, 1953, **35**, 119-129.
7. REED, J. D. Factors influencing rotary performance. *J. of Psychol.*, 1949, **28**, 65-92.

Department of Psychology
University of Rochester
Rochester 3, New York

DERIVATION OF A GENERALIZED CLASS OF IDEAL RETENTION CURVES BASED ON A CONCEPT OF EQUIPOTENTIALITY*

Department of Psychology, Brooklyn College

IVAN D. LONDON

In associating, as this writer has attempted (3-6), degree of retention with the proportion of hypothesized elements still participant in a neural complex that has been subject to a statistically defined process of decay, equipotentiality is implied. It matters not which element has become nonparticipant; what is important is that any given proportion of the original complex be thought as associated with a degree of retention, however indexed. This leads mathematically to the well-known exponential curve of forgetting, asymptotic to the abscissa of the plot, if no restorative process countering the progressive decay of the complex is assumed (3).

However, a further mathematical exploration of the consequences of equipotentiality leads to a generalized class of ideal retention curves in which the exponential curve turns out to be only one of several possible. The derivation to follow is based on a model which assumes varying degrees of equipotentiality, but excludes restorative considerations in order to keep the mathematics tractable and to underscore the essential idea.

Let us conceive of a minimum complex of neuron assemblies to be associated with initial 100 per cent retention on assumption of no overlearning. Since the complex is made the basis of retention, it will be referred to as an engram-complex to accord with earlier terminology. The neuron assembly itself will be designated as an engram-unit since it constitutes the building block of the complex. Subassemblies will correspondingly be called subunits.

Suppose, now, that each engram-unit contains n subunits¹ originally necessary for its formation and whose probability of remaining participant in the engram-unit after time t is e^{-at} , where a is taken as the "action coefficient." On the assumption of a degree of equipotentiality among these subunits, we may consider each engram-unit as remaining participant in the en-

*Received in the Editorial Office on October 4, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹A uniform n for each engram-unit is adopted for mathematical simplicity. In actuality, a nonuniform n must be presumed.

gram-complex as long as in each engram-unit the number of *subunits*, rendered nonparticipant with time, does not exceed r .

Instead of dealing directly with the proportion² of engram-units still participant after time t in the original engram-complex minimally associated with initial 100 per cent retention, we may allude more advantageously to their probability of remaining participant. We proceed now to calculate this probability under the assumptions stated above.

If the probability of participation for an engram-subunit after time t is e^{-at} , its probability of nonparticipation after time t is $1 - e^{-at}$. The successive terms of the binomial expansion,

$$[e^{-at} + (1 - e^{-at})]^n,$$

yield respectively the probability that after time t no subunit has been rendered nonparticipant, one subunit has been rendered nonparticipant, two subunits have been rendered nonparticipant, etc. (1, p. 106). The $(r + 1)^{\text{st}}$ term yields the probability that r subunits have become nonparticipant. The sum P of the first $(r + 1)$ terms then gives the probability that, after time t , an engram-unit continues participant in the engram-complex.

Plotting P against t generates "survival curves" which assign the probability of continued engram-unit participation in the original engram-complex after time t . These curves, then, are seen to be the sought-after ideal retention curves, and constitute a generalized class.

If we substitute x for e^{-at} , the first $(r + 1)$ terms of the binomial expansion, $[x + (1 - x)]^n$, is

$$[1] \quad P = \sum_{i=0}^r {}_nC_i x^i (1-x)^{n-i},$$

where

$${}_nC_i = n! / i!(n-i)!$$

It is highly desirable to obtain, if possible, an expression from which the summation sign has been eliminated. Close inspection, after appropriate manipulation, of the derivatives of [1] in the cases: $r = 1, 2, 3$ suggests that

$$[2] \quad \frac{dP}{dx} = (n-r) {}_nC_r x^{n-r-1} (1-x)^r,$$

where $0 < x \leq 1$ since $0 < t < \infty$. If it is recollected that

$${}_nC_r = (r+1) {}_nC_{r+1} / (n-r),$$

[2] may be readily proved by mathematical induction to hold in general.

Therefore, on integrating,

²As gauged by the savings method (3).

$$[3] \quad P = (n-r)_n C_r \int_0^x z^{n-r-1} (1-z)^r dz,$$

where z is substituted for x for purposes of integration in order to establish P as a function of the upper limit x .

By means of the incomplete β -function (1, p. 127, Ex. 25; 2, pp. 561-562), integration of [3] may be effected. Since

$$\beta_x(p, q) = \int_0^x z^{p-1} (1-z)^{q-1} dz,$$

$$[4] \quad P = (n-r)_n C_r \beta_x(n-r, r+1).$$

The coefficient of the incomplete β -function in [4] can be shown to be the reciprocal of the complete β -function. Since

$$\beta(p, q) = \int_0^1 z^{p-1} (1-z)^{q-1} dz = (p-1)!(q-1)!/(p+q-1)!,$$

$$\beta(n-r, r+1) = (n-r-1)!r!/n!.$$

But $(n-r)_n C_r = n!/(n-r-1)!r!$.

Therefore,

$$[5] \quad (n-r)_n C_r = 1/\beta(n-r, r+1).$$

Accordingly, on substitution of [5] in [4],

$$[6] \quad P = \beta_x(n-r, r+1)/\beta(n-r, r+1),$$

where $x = e^{-at}$.

The ratio of the incomplete to the complete β -function is known as the incomplete β -function ratio and has been tabulated (7). An appropriate transformation of [1] permits the ready calculation of this ratio also by

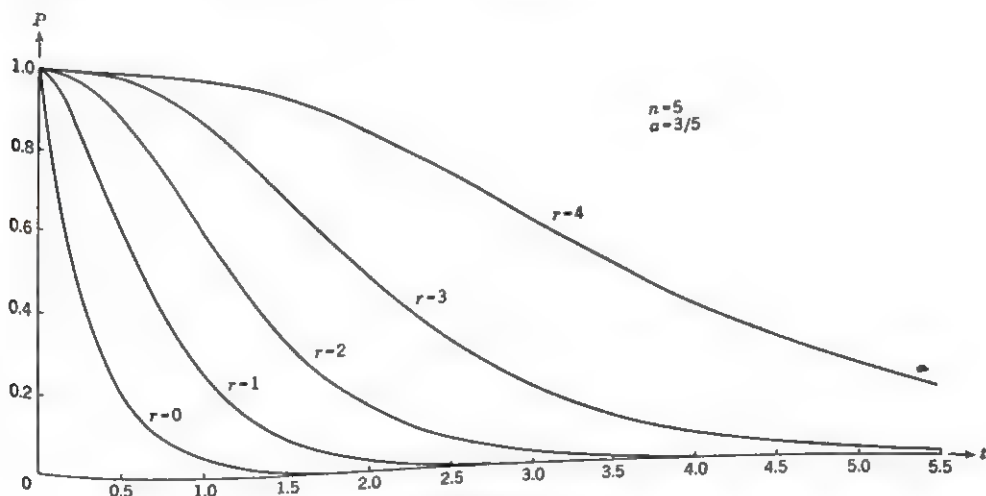


FIGURE 1

means of tables of cumulative binomial probabilities, recently made available (8). Illustrative plots of [6] have been made for the cases: $n = 5$; $r = 0, 1, 2, 3, 4$ with a taken as equal to $3/5$ in order to exhibit the nature of the general class of retention curves obtained on the assumption of degrees of intra-engram-unit equipotentiality and the absence of a restorative process to render previous engram-subunits again participant.³ For all n 's, $r = 0$ exponential curves result; for $n \geq 2$, $r \neq 0$, however, varieties of sigmoid curves are obtained.

With overlearning the associated engram-complex may be presumed to exceed the minimum complex postulated for the above derivation. The percentage excess of engram-units beyond the minimum constituting the latter engram-complex for $t = 0$ may then be utilized to determine the increase of probability of engram-unit participation for any given t as a result of overlearning. The resulting curves will accordingly show a continuation of P at the 100 per cent level, for t 's not in the neighborhood of zero, beyond that registered for the minimum complex, as well as a proportionate shift upward of the curves themselves with $P = 0$ still as asymptote. These upward-displaced curves add considerably to the coverage of the deduced class of ideal retention curves, again on the assumption of nonrestoration.

However, restorative processes, leading to reparticipation of engram-units and subunits in the engram-complex, should lead to asymptotic descent to $P \neq 0$, so that there should be produced a permanent leveling upward of the retention curves, originally derived on the assumption of nonrestoration. This has been mathematically demonstrated in a previous paper (3) for what now turns out to be the special case $n = 1$, $r = 0$ for retention associated with a minimum complex. In that paper it was shown that an asymptotic $P_\infty = 1 - a/b$ is a consequence of a restorative process countering one of progressive decay as indicated in the differential equation,

$$\frac{dP}{dt} = -aP + bP(1 - P),$$

³The magnitude of a may be graphically estimated from the empirical curve of retention by a method derived in a previous article (3, pp. 297-298); *non-overlearned* material of *minimal meaningfulness* being utilized to set up the corresponding curve of retention. The parameters n and r for a given empirical curve of retention may be estimated for non-overlearned material in general from calculations utilizing only three paired values of P and t , read off from the empirical curve in the non-asymptotic region of the curve; paired values of P and t , associated with levels of retention in the latter region, being inadvisable in practice due to the obscuring effect of residual asymptotic retention. This obscuring effect is attributable to restorative processes countering the progressive deterioration of the engram-complex (3) and results in degrees of deviation of the empirical from the ideal curve of retention. Details of the method, whereby the parameters n and r may be estimated from the empirical curve of retention, will be presented in a separate paper. The method makes essential use of Equation [2].

whose solution is

$$P = \frac{b - a}{b - ae^{-(b-a)t}}.$$

In general, it is not at all unreasonable to surmise that for cases of overlearning, particularly where meaningful relations are involved, a high ratio r/n should be anticipated, so that on assumption of restoration an asymptotic P_{∞} may be anticipated at a high level or even as indefinitely coincident with $P = 1$. These curves asymptotic to $P \neq 0$ again add further to the coverage of the generalized class of ideal retention curves previously derived, so that it is possible to view a great variety of disparate curves as products of essentially identical processes operating upon equipotentially diverse neural assemblages—a diversity essentially determined by the critical n 's and r 's involved.

REFERENCES

1. FELLER, W. *An Introduction to Probability Theory and Its Applications*. New York: Wiley, 1950.
2. FRANKLIN, P. *A Treatise on Advanced Calculus*. New York: Wiley, 1940.
3. LONDON, I. D. An ideal equation derived for a class of forgetting curves. *Psychol. Rev.*, 1950, **57**, 295-302.
4. ———. An ideal equation derived for overlearning. *Psychol. Rev.*, 1951, **58**, 54-59.
5. ———. A theoretical relation between rate of forgetting and internal body temperature. *J. Gen. Psychol.*, 1952, **46**, 151-157.
6. ———. An application of the engrammatic concept. *J. of Psychol.*, 1953, **36**, 379-385.
7. PEARSON, K. (Ed.). *Tables of the Incomplete Beta Function Ratio*. London: Univ. Coll. (*Biometrika* Office), 1934.
8. SIMON, L. E., & GRUBBS, F. E. (Eds.). *Tables of the Cumulative Binomial Probabilities*. Ordnance Corps Pamphlet: ORDP 20-1, 1952.

Department of Psychology
 Brooklyn College
 Bedford Avenue & Avenue H
 Brooklyn 10, New York

AUTISM, EMPHASIS, AND FIGURE-GROUND PERCEPTION*

University of Kansas and The Menninger Foundation

TEODORO AYLLON AND ROBERT SOMMER¹

A. INTRODUCTION

In the literature relating to need and perception, there are essentially two groups of studies. The first concern primarily the movement of the cognitive processes in the direction of need-satisfaction. The experiments of Schafer and Murphy (9), Proshansky and Murphy (7), and Jackson (4), may be regarded as prototypes of this approach. On the obverse side are those dealing with perceptual defense. These latter, seemingly innumerable, usually posit a typology of reaction patterns ranging from defense (raising of threshold to punished stimuli, similar to autism in this sense) to vigilance (lowering of threshold or sensitization to punished stimuli).

The roots of this concept of vigilance go deep. One may remember the classic experiment by Tolman, Hall, and Bretnall (11) in which the alerting or emphasizing effect of electric shock was clearly evident. Many similar investigations followed at its heels; Jones (5) provides a good bibliography. Two recent studies by Pustell (8) and Dulany (3) are pertinent. Pustell approaches the concepts of vigilance and defense using a cue-drive hypothesis. He maintains that if escape from punishment is impossible and anxiety is very intense, defense will appear. When such anxiety is more moderate in intensity, a vigilant attitude will prevail. Dulany employs a functionalist approach to the learning of the perceptual patterns of defense and vigilance. He notes that responses instrumental to avoidance of punishment will be strengthened while those associated with punishment will become weaker.

The present study, stemming from the line of research begun with the City

*Recommended for publication by Gardner Murphy, received in the Editorial Office on October 14, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹The present study is part of a larger program of perceptual learning research under the direction of Dr. Gardner Murphy. This investigation was supported by a research grant (M-715C) from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service. In carrying out the investigation, we gratefully acknowledge the assistance of Miss Diana Laskin, Mr. Seymour Liss, and Mr. Cyril Feldstein. The study is described in greater detail in the thesis by the senior author submitted in partial fulfillment of the degree of Master of Arts to the Graduate Faculty, University of Kansas.

College investigations on the effects of need on perception, has a somewhat different orientation. Yet our hope is also to tie together the findings from both the autism and the emphasis studies in terms of the situations in which defense or alerting will appear. It is based on a pilot study by the writers in which electric shock was utilized in a tactual-kinesthetic figure-ground situation (to be described later in this paper). In brief, it was found that there was a slight tendency for those Ss who reported experiencing the shock as *slightly* unpleasant to respond to the non-punished aspect in the figure-ground situation (as would be predicted from autism or defense theory); and a marked tendency for those Ss who reported experiencing the shock as *moderately* or *very* unpleasant to respond to the punished alternative (which would be predicted on the basis of the concepts of alerting, emphasis, or vigilance).

It was felt that if this finding could be confirmed, it would have important implications for need-perception theory in that defense and vigilance could be conceived of as functional behavior patterns, adopted by the individual when a particular type of situation arose. It might also reconcile autism theory to some degree with much of the critical discussion of the "new look" studies that has appeared in the journals. The example of the skier's attending to the (punishing) treestumps along the trail could be considered a case of considerable (potential) punishment that would call for alerting to the sources of danger. Focussing on the stumps would assist in making a smooth run down the trail more than would the ignoring of them. There is likewise the case of the hysteric obsessed with the fear of snakes. Rather than overlooking a snake in a field, she tends to accentuate any rustling of the grass, and to reshape every available branch or piece of wood in the field. It should be apparent that an approach envisaging an invariant minimization of unpleasantness is not adequate to explain such data. Some method is needed whereby those cases in which alerting will occur can be distinguished from those in which defense will appear. The present paper proposes that one such method entails the use of degree of hedonic tone involved. Different predictions are made according to S's experience of pleasantness-unpleasantness. The stimulus situation used happens to fall under the rubric of tactual-kinesthetic figure-ground organization. There is no particular reason why the hypotheses could not have been tested in the visual or auditory modalities. From this standpoint, the preceding cases of the skier and the hysteric can be considered examples of an individual attending to sources of extreme unpleasantness. From the hypotheses to be outlined, we would predict that if the discomfort had been *minor*, there would have been movement of the

cognitive processes in the direction of rewarded or non-punishing aspects of the field.

Specifically, the hypotheses to be tested in the present study, in which electric shock is employed in a tactual-kinesthetic figure-ground situation, are as follows: (a) If the subject experiences the electric shock as *slightly* unpleasant, he will tend to perceive the *non-punished* aspects of the field. (b) If the subject experiences the shock as *moderately* or *very* unpleasant, he will tend to perceive the *punished* aspects of the field.

Experience of the shock as to degree of unpleasantness will be determined in two ways. The first method entails S's rating his reaction along a five-point scale; the second involves E's rating of S's behavioral reaction to the shock. It should be noted that neither method is directly related to the amount of shock used in the study. That is, the severity of shock is defined phenomenally and behaviorally rather than physically in terms of the voltage or amperage of the inductorium.

B. APPARATUS

The stimulus materials used in the present study consisted of a set of two 3-dimensional reversible profiles somewhat similar to those used in the Schafer-Murphy investigation (see Figure 1). The profiles were formed by grooves in 12" x 12" plaster casts. The grooves were 8" in length, $\frac{5}{8}$ " in width, and $\frac{1}{8}$ " deep. There were four of these plaques: Rufus, a left-pointing face; Clem, a right-pointing face; Horace, a full face used as a set-breaking figure; and a post-test profile line similar to both Rufus and Clem. As the right-pointing, left-pointing, and ambiguous post-test profiles were cast from the same mold, the center profile line was identical for each figure.² Meaningful backgrounds were painted in colored enamel on the plaques to preclude, for Rufus and Clem, visual perception of the other face during the tracing.

A rubber glove was worn on the tracing hand to minimize extraneous tactual cues and compel S to rely solely on his figure-ground organization of the center line in determining the identity of the particular face.

A Lab-Tronics Physio-Stimulator (Model N-103) set at 45 volts was used with the first 24 Ss. When repairs on the apparatus became necessary, a Variac type UCIM inductorium which appeared to give a comparable

²In the pilot study alluded to previously, this identity of profile lines for the three figures was established in that Ss, when blindfolded, were unable to distinguish the actual Rufus from the actual Clem plaques. On two trials of attempting to identify each of the two figures, only one of the 22 Ss correctly identified the plaques. This is below chance expectancy.

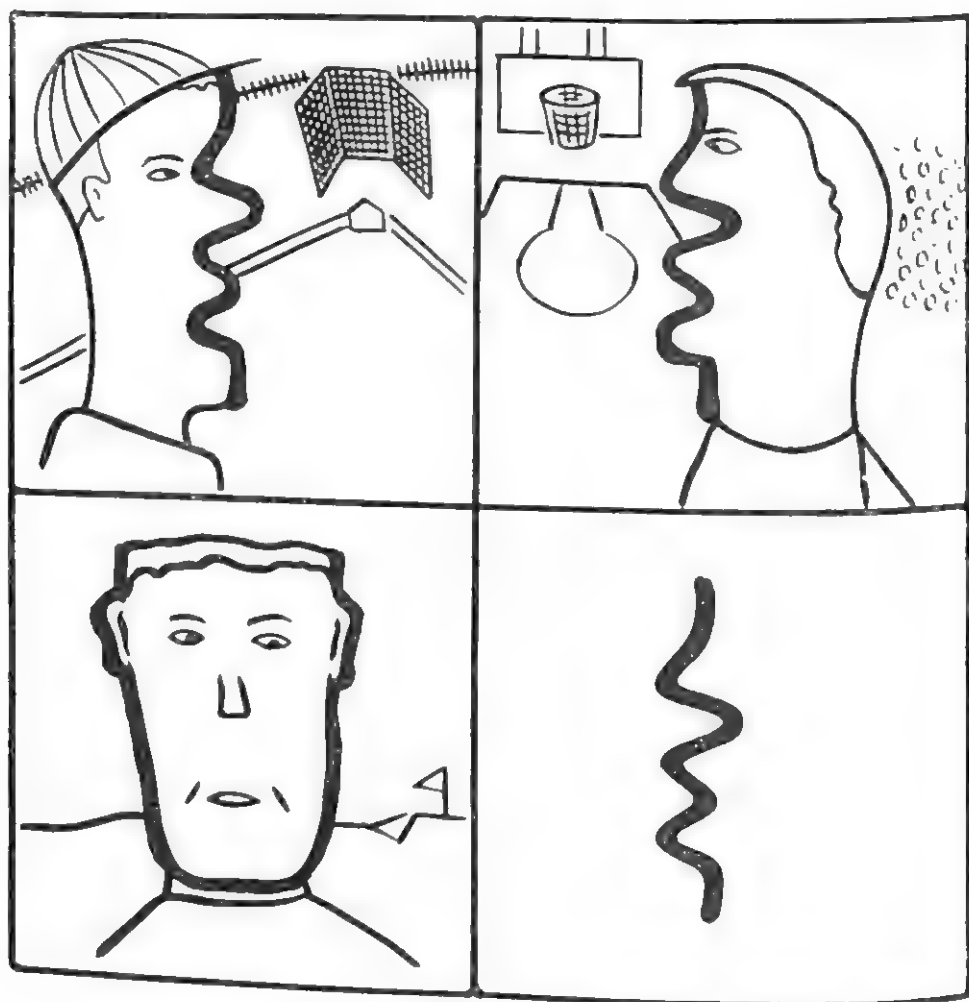


FIGURE 1
CLEM—RUFUS
HORACE—AMBIGUOUS FIGURE

shock was used with the remaining six *Ss*. Analysis of the data showed no change in results for these *Ss*. The electrodes were mounted on a small piece of wood covered by a "finger glove" in which *S* inserted his index and middle fingers.

A black lightproof piece of cloth was used as a blindfold in the post-training series.

C. PROCEDURE

Ss were 30 female undergraduate students enrolled in the elementary psychology course at the University of Kansas. None of these *Ss* knew

prior to the experiment that shock would be used. Ss were tested individually in a small experimental room, seated across a table from E. Upon entering the room, S was told:

Hello. Would you please sit down? I am going to show you a series of profile grooves. You see this plaque? What you are expected to do is trace the grooves embedded in it like this. (E traces the groove with his index finger and then puts the plaque away.) Here is another plaque and you are to do the same with this one. (E traces the groove and then puts the plaque away.) I have a third plaque; you do the same with this one too. (E traces the groove and then puts the plaque away.) Now each of these plaques has a name. Every time I present the plaque to you I will call out its name, and you will proceed to trace it twice. It is expected that you will learn these faces so that you will be able to identify them when blindfolded. Is that clear? OK, now see if you can wear this glove on your right hand. Since you will be using your index finger this should be as tight as possible. All right, now I want you to put your left two fingers into this shock apparatus.

E presented one face at a time and called out the appropriate name while putting the face in front of S. The order of presentation of the faces was as follows: Rufus, Horace, Clem, R*, H, C, C**, R, H, R*, C**, C, H, H, R, R*, C**, R*, C**, H, R, R*, C**, H.

Some of these trials were followed by electric shock. These are indicated by single asterisks for Ss for whom Rufus was to be shocked and double asterisks for Ss for whom Clem was to be shocked. Shock was associated with Rufus for 17 Ss, and with Clem for the remaining 13. The shock came three seconds after the tracing to avoid the criticism that the shock interfered with the tracing process itself. Also it can be noted that there were *in toto* only five shocks for any given S. The exact sequence of events for the plaque associated with shock (punished plaque) was always as follows: beginning of tracing; end of tracing; three-second waiting period; shock of one second duration; plaque put away; E presents next plaque. For the plaque not associated with shock, the sequence was as follows: beginning of tracing; end of tracing; three-second waiting period; plaque put away; E presents next plaque.

After the last plaque was presented E told the S:

All right; you can take your fingers off the shock apparatus. Now, I'm going to blindfold you and we'll see if you have learned the faces. (E blindfolds S.) I shall start you off on each face and whenever you think you have identified the face, call out the name. If you are in doubt you can trace the same face twice.

In the post-training series there was a total of 27 presentations: 14 of

the ambiguous profile line, and 13 of Horace (set-breaking full face). The latter was alternated with the ambiguous profile in the following sequence. Ambiguous, Horace, H, H, A, A, H, A, H, H, A, H, A, H, A, A, H, A, H, H, A, H, A, A, H, A, A.

While the *S* was tracing the first plaque in the post-training series (*before* her response) *E* rated her in terms of the overt reaction she had shown to the shock: strong, moderate, and slight. The criteria used by *E* for rating *S*'s reaction to the shock as *strong* were such items as flinging of the arm whenever shock was given, raising of finger from the table in absence of shock, and verbalizations regarding the shock (e.g., "Gee, that hurts," "How long is *this* going to last?"). If the *S* showed any or all of these characteristics to a lesser degree, *E* rated her reaction to the shock as *moderate*. On the other hand, if *S* showed few or none of the above mentioned behaviors, *E* rated her reaction to the shock as *slight*.

After the last plaque was put away, *E* proceeded to remove *S*'s blindfold and assisted her in taking off her glove. This was followed by a prepared interview as to reaction to shock, awareness of face-shock relationship, confidence in judgments in the post-test series, face remembered most vividly, and physiognomic qualities of the faces. For these items, *S* was requested to indicate her feelings by checking some point along an appropriate continuum (e.g., when rating affective value of the shock, she was asked to check along a scale ranging from "not unpleasant" to "unbearable.")

At the conclusion of the session, *E* apologized to *S* for having used the shock and requested her not to mention the study to any of the other students.

D. RESULTS

Prior to the study, it was decided to treat the results in two ways. One, in terms of the first response to the ambiguous plaque. Hochberg³ and others have stressed that in an ambiguous figure-ground situation, it is the first response that is the most meaningful; i.e., least contaminated in terms of subsequent responses being determined by many extraneous factors, the most important being the desire to perceive all the alternatives that had been presented in the pre-test series. For example, many subjects would become puzzled after they had given several Rufus responses and inquire as to the whereabouts of Clem. The second method of treating the data is in terms of *S*'s responses to all 14 presentations of the ambiguous plaque. The *S*s can also be classified by the two criteria alluded to previously; namely their

³Hochberg, J. E. Personal communication.

own rating on the scale from not unpleasant to unbearable; and *E*'s rating derived from overt manifestations of anxiety.

Analyses of the *Ss*' first response to the ambiguous figure failed to reach significance when *Ss* were classified by their own rating and also when the *Ss* were classified by *E*'s rating ($p = .08$ in predicted direction).

Table 1 is based on all 14 responses by each *S* to the ambiguous plaque. *Ss* are classified by their own rating of reaction to the shock. As predicted

TABLE 1
TOTAL RESPONSES TO AMBIGUOUS FIGURE WITH *S* CLASSIFIED BY OWN RATING

	n	Mean number of responses to punished figure	Sigma	<i>t</i>	df	p (one-tail)
<i>Ss</i> reporting shock as slightly unpleasant	10	5.60	2.73	2.78	28	.01
<i>Ss</i> reporting shock as moderately or very unpleasant	20	8.85	3.23			

by the hypotheses, the group which had rated the shock as moderately or very unpleasant reported perceiving more punished faces in the test series than the group which had rated the shock as slightly unpleasant. The difference between the groups is significant at the .01 level.

Table 2 resembles Table 1 except that *Ss* are classified by *E*'s rating rather than their own report. It presents data based on all 14 responses

TABLE 2
TOTAL RESPONSES TO AMBIGUOUS FIGURE WITH *S* CLASSIFIED BY *E*'S RATING

	n	Mean number of responses to punished figure	Sigma	<i>t</i>	df	p (one-tail)
<i>Ss</i> reporting shock as slightly unpleasant	10	7.80	3.27	—	27	NS
<i>Ss</i> reporting shock as moderately or very unpleasant	19	7.97	3.56			

to the ambiguous profile for both the slightly unpleasant and the moderately and very unpleasant groups. The picture here is quite different from that found previously. When classified on the basis of overt reactions to shock, the predicted relationships do not appear.

A compromise solution to the problem of determining the most valid in-

dex of *S*'s experience of the shock is arrived at by examining only those cases in which *S*'s own rating of experience agrees with *E*'s rating. That is, cases where, for example, *S* rated the unpleasantness as moderate and *E* rated her reaction as slight or strong, would be excluded. The results from this analysis and based on the *first response* to the ambiguous figure are in the predicted direction and show 12 cases correctly predicted to 5 incorrectly predicted ($p = .07$, one-tail). The results based on *all 14 responses* to the ambiguous figure show a difference between the slightly and moderately-very groups that is significant beyond the .05 level in the predicted direction.

During the interview, after *Ss* had been requested to rate their feeling toward the shock, they were asked which of the faces they remembered most vividly. Then they were asked which face they remembered next most vividly. The reason for this latter question was that the vast majority remembered Horace, the set-breaking full face, most vividly due to its distinctiveness. The only two *Ss* who answered the first question with a face other than Horace had both rated the shock as moderately to very unpleasant and both gave the punished face as the most vivid. Table 3 presents the

TABLE 3
FACE (EXCLUDING HORACE) REMEMBERED MOST VIVIDLY

	Remembered most vividly		p
	Punished face	Non-punished face	
<i>Ss</i> reporting shock as slightly unpleasant	2	8	.02
<i>Ss</i> reporting shock as moderately or very unpleasant	13	6	

face other than Horace mentioned as remembered most vividly. The trend is quite clear and is significant at the .02 level in the expected direction. Those *Ss* who rated the shock as slightly unpleasant tended to remember the non-punished face, while those *Ss* who rated the shock as moderately or very unpleasant tended to remember the punished face.

Relationships between the other variables (awareness of face-shock relationship, physiognomic characteristics of the faces, and confidence in judgment in the ambiguous situation) are treated in detail elsewhere (1).

The foregoing tables establish relatively clearly that the "slightly unpleasant" group derives from a different universe from the "moderately-very unpleasant" group in the matter of post-test responses and recall of faces.

The differences between the groups are in the direction predicted by the hypotheses. However, another method of treating the data is to check the significance of the trends within the groups themselves. This treatment more closely approaches the original hypotheses which were phrased in terms of intra-group differences. Table 4 summarizes these comparisons of signifi-

TABLE 4
INTRA-GROUP COMPARISONS OF PUNISHED AND NON-PUNISHED RESPONSES

	n	Mean number Punished Responses	Mean number Non-punished Responses	t	df	p (one-tail)
Ss reporting shock as slightly unpleasant	10	5.60	8.40	1.54	9	.07
Ss reporting shock as moderately or very unpleasant	20	8.85	5.15	2.64	19	.01

cances within the groups (Ss classified by own rating) for all 14 responses to the ambiguous figure. A *t*-test on the mean of the differences between rewarded and punished responses was computed for each of the two groups. It can be seen that for the moderately-very group, the number of punished responses exceeds the number of non-punished responses and this is significant at the .01 level. For the slightly group, the number of non-punished responses exceeds the number of rewarded responses and this is non-significant, though it reached the .07 level.

Parenthetically it can be noted that there was a strong response preference for Rufus, the left-pointing face (277 Rufus to 143 Clem responses). This also appeared in the pilot study (177 Rufus to 129 Clem responses), and can possibly be attributed to a (supposed) preference of Ss from a predominantly right-handed population for left-pointing profiles.

E. DISCUSSION

The present study should have interesting implications for need-perception research. To begin with, it presents testimony for the necessity of adequate calibration of the affective value of the pleasant or unpleasant stimuli. There is a point along the unpleasantness continuum where an alerting rather than a defensive trend appears. If the experimenter is not careful, he may be employing stimuli both to the right and left of this point and thereby confound his results.

Furthermore, the experiments of Smith and Hochberg and of Pustell suggest that there is a third inflection point along the unpleasantness con-

tinuum, beyond which defensive behavior will again prevail. This point may be designated as the "unbearable" position on the present hedonic scale. At this point we may suppose that the unpleasant stimulus is too unpleasant for *S* to contemplate; and hence avoidance behavior makes its appearance. This piece of conjecture derived from an attempt to reconcile our results with those of Smith and Hochberg (10), who found that with the Schafer-Murphy reversible faces in the visual modality, the non-shocked alternative tended to be reported more frequently.

To test the validity of this hypothesis in a preliminary fashion, several *Ss*, not belonging to the 30 *Ss* tested in the experiment, were tested using a markedly stronger intensity of shock (65 volts with electrode jelly). *Ss'* reactions and verbalizations supported our feeling that the desired degree of unpleasantness was produced. Were the results as we had predicted? In the main, yes, but there is some equivocality due to the reluctance of several *Ss* to use the extreme end of the graphic continuum. For example, when *E* noted that an *S* who had continually been in a state of extreme tension checked below the "very unpleasant" point, he felt it necessary to inquire further as to the intensity of her reaction. Upon questioning, *S* readily acknowledged that she had experienced the shock as unbearable. When asked directly to "rate it as accurately as you can," she changed her check to slightly less than the unbearable point. However, the danger of suggestion by *E* (who had at this point heard *Ss* responses notwithstanding efforts to refrain from acting on this knowledge) remains. Thus, although the results cannot be accepted as valid because of difficulties arising in regard to *S's* scale rating, there does seem to be a point on the hedonic scale lying somewhere between very unpleasant and unbearable, beyond which *S* cannot tolerate the punished face and must display defensive modes of behavior. Emphasis on the punished face disappears and is replaced by an autistic repression of the source of such extreme unpleasantness. There might also be the disruption effect Tolman speaks of.

On this basis, we may engage in some conjecture as to the reaction patterns corresponding to the breaks along the hedonic scale. To the right of the first break are found those *Ss* who have an experience of slight unpleasantness. Autism theory would predict that these *Ss* would tend to minimize, ignore, or defend against such slight unpleasantness, perceiving in the direction of rewarded or non-punished aspects of the field. The second group contains those *Ss* who experience the unpleasantness as moderate or strong (there is some lack of precision here in that *Ss* interpretations of the scale values are not always comparable). The present study found these *Ss* em-

phasizing the punished aspects of the field. The trend was quite unmistakable; our *Ss* tended to respond to the punished face in the post-test series and remember it most vividly after the testing was concluded. Finally, there may be a third point along the continuum where the unpleasant stimulus is just too noxious for *S* even to contemplate, so she again engages in defensive or repressive behavior. The results found by Smith and Hochberg are quite intelligible on this basis, as are those secured by the present writers after the present investigation had been concluded. A *U*-curve theory of effect of affectivity in terms of alerting to sources of negative affect seems to order most of the data quite sensibly. Thus with slight unpleasantness, defensive behavior may appear; with moderate unpleasantness, alerting or emphasis; and with intense unpleasantness, again defense or repression.

However, the problem of explaining these results still remains. It is quite apparent that the "slightly unpleasant" group differs from the "moderately-very" group in their post-test responses. One explanation is in terms of the scale checks of "slightly," "moderately," or "very unpleasant" themselves reflecting underlying personality structures which determine either autistic or vigilant modes of perceiving. The intensity check marks can be considered dependent variables, highly correlated with particular personality predispositions. It is known that females tend to react more strongly to shock than males. There is probably also a positive correlation between hysterical tendencies and reaction to shock. If this were the case, our results might then indicate that hysterics alert themselves to sources of unpleasantness, while individuals of another personality disposition tend to minimize the unpleasantness of the shock.

Our train has thus pulled in to the very station we had hoped to avoid. Experience of intensity of shock was such a deceptively simple variable that we had neglected to attend to the individual experiencing the shock. A needed step is to determine whether each individual has a baseline of reaction to shock wherein mild shock engenders autism while strong shock produces vigilance or whether the check marks themselves reflect underlying personality structures. The present study was set along these latter lines but has left unanswered the question of the actual significance of the check marks themselves. The alternative explanation involves a functional view of defense and vigilance. That is, the same individual will adopt one, then the other, depending on the intensity of the experience of the shock. On the basis of our results, we should expect defense to be used with mild shock and alerting with strong shock.

Thus the former explanation refers to "people who" predominantly use

defense and vigilance, the experience of intensity of shock being at most a dependent variable. Individuals using defense are those with personality dispositions to under-react to electric shock while those who habitually use vigilance are those who over-react. This explanation requires one to predict that if the intensity of experience of shock were varied for the individual, he would still adhere to his previous defensive or vigilant mode of perceiving.

Another interpretation of the present data is in terms of the concept of distraction. It could conceivably be maintained that the shock or, more precisely, fear of shock, distracted *S* from learning the profile line of the shocked face. This conclusion does not seem warranted for several reasons. First, the results themselves, for those *Ss* who experienced the shock as moderately or very unpleasant (who should have been subjected to even greater distraction than the slightly unpleasant *Ss*), show a significant tendency to emphasize the punished face. It would be a peculiar type of distraction that facilitated the learning of the pattern that it was interfering with. No doubt there is considerable distraction value attached to the shock as there would be to any reward or punishment used in the experiment, but it is not the shock as a distractor from *learning* the faces that produced the present results (especially in terms of the difference between the "slightly" and "moderately-very" groups which seems incompatible in direction with what would be predicted on the basis of a distraction hypothesis). Also, the shock was administered three seconds *after* *S* had finished tracing the profile line and there were only 5 ("objectively" rather mild) shocks delivered during the entire procedure. It is not denied that the shock, or any other closely contiguous stimulus for that matter, had distraction value for the subject, but only that on the basis of both procedure and results, an interpretation of the rôle of the shock in terms of emphasis and autism seems far more tenable.

One final note concerns the sense modalities used by the subjects during the procedure. The hypothesis that *S* mainly used the tactual modality during the tracing of the profile, soon becomes inadequate when one realizes that *S* was provided with rubber gloves to minimize tactile cues. The feeling of the writers is that *Ss* tend to use a kinesthetic approach to learning the profiles, i.e., Rufus goes "in and out, in, out and down"; Clem goes "out, in, out, in and down"; Horace goes "down, way around, and up." If we were to pursue the point further, it would soon lead into the nature of tactual perception and finally to the question of whether all of this perception is ultimately visual in terms of the imagery which *S* actually uses. Also

it would be difficult to avoid becoming involved in a distinction between perception and cognition, as it could easily be maintained that *S* perceives the profile groove as merely a wavy line, and subsequently cognized it into a right- or left-pointing face. It is our preference to use a definition of perception similar to that of Coutu (2), in terms of giving a stimulus a meaning, wherein the present cognitive attribution of right-pointedness or left-pointedness would be subsumed under the generic rubric, perception. From this standpoint, the present stimulus situation has been referred to as involving tactual-kinesthetic figure-ground perception. Whether a better term exists, is an open question, yet it is precisely the region of shadows and chimeras that the cartographers of the reality of man may find the most fruitful for exploration.

F. SUMMARY

The present study represented an attempt to confirm previous findings from a pilot study in which *Ss* who rated the electric shock (used as punishment) as "moderately" or "very unpleasant" tended to perceive the punished aspect of a tactual-kinesthetic figure-ground situation. It also aimed at confirming a slight tendency of *Ss* who rated the shock as "slightly unpleasant" to perceive the non-punished alternative.

Specifically, the hypotheses to be tested were as follows: (a) If *S* rates the shock as *slightly* unpleasant, he will tend to perceive the *non-punished* aspects of the field. (b) If *S* rates the shock as *moderately* or *very* unpleasant, he will tend to perceive the *punished* aspects of the field.

Thirty college girls were used as *Ss*. The experiment was divided into three parts: training series, post-training series, and interview.

In the training series, *Ss* traced three face outlines with the index finger. Two of these were profiles with an identical contour line, the third a set-breaking full face. A periodic reinforcement of electric shock was used to associate punishment with a given profile.

During the post-training series, *S* was blindfolded and asked to identify the ambiguous contour line (identical in the right- and left-pointing profiles). The results supported the hypotheses; namely, *Ss* who rated the shock as moderately or very unpleasant tended to report the *punished* face. On the other hand, *Ss* who rated the shock as slightly unpleasant tended to report the *non-punished* alternative.

Following the model of a U-curve theory of unpleasantness, implications for defense-vigilance, autism-emphasis theory were sketched, revolving about the severity of the unpleasantness to which *S* is subjected as determining whether he will perceive punished or non-punished aspects of the field.

REFERENCES

1. AYLLON, T. The effect of punishment on tactual-kinesthetic figure-ground perception. Unpublished master's thesis, Univ. Kansas, 1955.
2. COUTU, W. Emergent Human Nature. New York: Knopf, 1949.
3. DULANY, D. E. Avoidance learning of perceptual defense and vigilance. Unpublished doctor's dissertation, Univ. Michigan, 1955.
4. JACKSON, D. N. A further examination of the rôle of autism in a visual figure-ground relationship. *J. of Psychol.*, 1954, **38**, 339-357.
5. JONES, H. E. Trial and error learning with differential cues. *J. Exper. Psychol.*, 1945, **35**, 31-45.
6. MURPHY, G. Affect and perceptual learning. *Psychol. Rev.*, (in press).
7. PROSHANSKY, H., & MURPHY, G. The effects of reward and punishment on perception. *J. of Psychol.*, 1942, **13**, 295-305.
8. PUSTELL, T. Cue and drive aspects of anxiety in relation to perceptual vigilance and defense. Unpublished doctor's dissertation, Univ. Michigan, 1955.
9. SCHAFER, R., & MURPHY, G. The rôle of autism in a visual figure-ground relationship. *J. Exper. Psychol.*, 1943, **32**, 335-343.
10. SMITH, E. E. P., & HOCHBERG, J. The effect of "punishment" (electric shock) on figure-ground perception. *J. of Psychol.*, 1954, **38**, 83-87.
11. TOLMAN, E. C., HALL, C. S., & BRETNAIL, E. P. A disproof of the law of effect and a substitution of the laws of emphasis, motivation, and disruption. *J. Exper. Psychol.*, 1932, **15**, 601-614.

Department of Research
The Menninger Foundation
Topeka, Kansas



THE EFFECTS OF MONETARY REWARD AND PUNISHMENT ON AUDITORY PERCEPTION*

The Menninger Foundation

FRED W. SNYDER AND CHARLES W. SNYDER¹

A. PROBLEM

Previous studies of the effects of arbitrary reward and punishment upon response to ambiguous figure-ground situations have been limited to the visual modality. Schafer and Murphy (3) consistently rewarded one half-moon face and punished another by means of giving and taking money from *S*. In the test series the two faces were combined into a reversible figure-ground pattern. The *Ss* reported predominantly the "rewarded" face. A repetition, with modifications, by Rock and Fleck (2) did not confirm these results. However, Jackson (1) repeated both the Rock-Fleck and Schafer-Murphy experiments. When he followed the procedure of Rock and Fleck, results were similar to their findings. When the Schafer and Murphy procedure was followed, results similar to theirs were found. A further test by Smith and Hochberg (4), using electric shock for punishment, also tended to confirm the Schafer and Murphy results.

The hypothesis of the present study was that perception of an auditory figure-ground situation can be modified by reward and punishment. More specifically, we predicted that a significant number of *Ss* would give a greater number of verbal responses for the "rewarded" voice than for the "punished" voice.

B. METHOD

1. Design

The Schafer-Murphy (3) design was modified for the auditory modality. The *S* was presented two voices speaking at the same time and reported at the end of each trial as much of the content as possible. This test was preceded by a training series of the same voices speaking separately in which *Ss*

*Recommended for publication by Gardner Murphy, received in the Editorial Office on October 14, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹The authors wish to thank Dr. Gardner Murphy for his guidance throughout. We wish to thank Donald P. Spence, James E. Simpson, and Robert Sommer for their help in the design. This investigation was supported by a research grant (M-715C) from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service.

winning or losing of nickels was paired with particular voices. No overt response was required of *S* during the training.

Scores for *S* were obtained from the words reported of each voice. Whether *S* had reported predominantly words of the rewarded or punished voice was then determined by comparison of his test scores. A sign test was employed to determine the significance of the number of *Ss* supporting the hypothesis. A *t*-test of difference scores was used to determine if the combined scores for the rewarded voice were significantly greater than for the punished voice.

2. Apparatus

The material to be read was selected qualitatively for its ease in comprehension, interest to *S*, and relative neutrality with respect to pleasantness-unpleasantness. Eighteen sentences from Rachel Carson's *The Sea Around Us* were randomly selected for the training. A different random order of the same sentences was used in the test.

For the training, a Magnecorder tape was made of two adult males, "D" and "J," speaking separately. Assignment of voices to the training sentences was random with a voice restricted to not more than two consecutive sentences. For the test period a similar recording was made of both "D" and "J" speaking into the microphone at the same time without regard for each other's voices. An approximate balance in the speech qualities of the two voices was obtained. The test tape was cut into six trials of approximately 10 seconds each.

3. Subjects

The 41 *Ss* (17 males; 24 females) of this experiment were obtained from a local business school in two samples of 20 and 21 *Ss*. An interval of three months elapsed between use of the two samples. The mean age of the combined sample was 19.8 years with a range of 17 to 26 years. In addition to the money he won during the training series, each *S* was paid one dollar per hour. The *Ss* were randomly assigned to one of the two reward conditions: Reward Voice "D"—Punish Voice "J," or Reward Voice "J"—Punish Voice "D."

4. Procedure

The *Ss* were brought individually to the experimental room and seated at a table. The tape recorder on the table was set to reproduce the voices at an average sound level of 62 decibels (measured with General Radio Sound-Level Meter, Type 1551-A). Each *S* was then given the following instructions:

This is a procedure in which I'm going to play the tape recorder; also I have some money here. I'm going to stack some up and that's going to be yours. I'll put some money here and that will be mine. We're going to play a game with the money. You may get all of the money or you may wind up with none of it, or something in between. What's in your pile you may take with you when you go. You don't have to say anything, but you do have two jobs to do. You have to listen to what's on the tape and you also have to watch whether you're getting money or if it's taken away. Do you have any questions before we start?

The training was begun with two piles each of 10 nickels, one for *E* and one for *S*. During the playing of each voice, *E* transferred nickels between his pile and *S*'s, the direction of the exchange being dependent upon which voice was to be rewarded for the particular *S*. In order for *S* to complete the training series with more money than when he had begun, the transfer of nickels had to be unequal. The training series was played through twice. During the first playing, from one to three nickels were exchanged in each transaction in such a manner that *S* appeared to be losing, although by the end of the run he appeared to be breaking even. During the second playing the transfer of coins now favored *S* and he was 80 cents ahead by the end. This procedure was adopted in order to sustain the interest of *S*. To minimize the possibility that *S* would discover that the voices were the cue for the exchange, *E* waited a variable time after the beginning of each trial before transferring money and completed the transfer before the voice ceased. At the completion of the training, which took four and one-half minutes, *S* was told that this was the end of the first part of the procedure and reminded that the money in his pile belonged to him and was to be taken when he left.

Then *E* said:

I'm going to play a short portion of the tape. Listen and try to get as much as you can out of it. When I stop the tape, repeat word for word, as much as possible, what you heard. Do you have any questions?

The same two voices, speaking the same material as in the training, were then presented together in competition. Although there were normal speech pauses during which only one or neither voice was speaking, most of the time the voices were intermixed. At the end of each trial, *S*'s reproduction of the taped presentation was recorded on a Dictaphone. After the first trial, *E* said, "All right, what did you hear?" When *S* had responded *E* said, "Fine," regardless of the voice that was repeated. After the other trials he said merely, "All right."

After the last trial, *S* was asked whether he listened during the test to both voices, or to one more than the other and asked to identify the voice. Information regarding *S*'s relative preference for the voices was obtained as well as other bases for discrimination between the two. He was asked if he knew why he was receiving and losing money and if he were aware of the purpose of the experiment. He was also asked if anyone had told him about the experiment before coming, and was cautioned not to discuss the procedure with other *Ss*.

C. RESULTS

Each *S*'s responses were scored on the basis of the following criteria: *4 points*—Three or more words in succession in one sentence repeated exactly (counting all words, including articles, conjunctions, etc.). *3 points*—Three or more words in succession in one sentence repeated but allowing synonyms to count. *2 points*—Three or more words in one sentence or their synonyms but words not in proper sequence. *1 point*—One or two words or their synonyms in one sentence repeated but words not necessarily in proper sequence.

The scoring unit was one sentence in a particular voice regardless of shifts to the other voice and back again. These scoring criteria were used in place of a simple word count in order to retain the characteristic of grouping of words into meaningful units. Thus, more weight was given to meaningful groups and less weight to isolated words. In addition, there was a cut-off point at the maximum possible score for each sentence. This minimized the factor of differential ability in recall among *Ss*.

The scoring technique yielded two scores for each *S*, one for the "rewarded" and one for the "punished" voice in each trial. For the analysis, the scores for the six trials in the test were totaled. If more of the response units were those associated with the rewarded voice, the *S* was said to add support to the hypothesis.

The central finding was that 31 out of 41 *Ss* responded predominantly to the previously rewarded voice, the sign test yielding $\chi^2 = 9.76$; 1 *df*; $p < .01$. An additional statistic, one which tested the effects of the two training conditions (Table 1), yielded $\chi^2 = 9.47$; 1 *df*; $p < .01$. A sign test of the difference between voices (Table 2) resulted in $\chi^2 = 3.64$; 1 *df*; $p = .06$. To test further the difference between reward conditions, an analysis of variance (Table 3) was made, using as measures the differences between test scores for the "rewarded" and "punished" voices. Because of the lapse of time between the first 20 and the last 21 *Ss*, a test of the difference between these samples was included in the analysis of variance. Re-

sults of all Ss were combined since the difference between voices, between samples, and the interaction of voices and samples were not significant. A t -test of the difference scores was made, resulting in $t = 2.61$; 40 df ; $p < .02$.

TABLE 1
DISTRIBUTION OF SUBJECTS IN RELATION TO VOICE PREDOMINANTLY REPORTED IN TEST AND TO TRAINING CONDITION

Voice predominant in test result	Training condition		Total
	"D" rewarded "J" punished	"J" rewarded "D" punished	
"D"	12	2	14
"J"	8	19	27
Total	20	21	41

TABLE 2
DIFFERENCES BETWEEN TRAINING CONDITIONS IN TERMS OF THE NUMBER OF SUBJECTS SUPPORTING AND NOT SUPPORTING THE HYPOTHESIS

Test result	Training condition		Total
	"D" rewarded "J" punished	"J" rewarded "D" punished	
+	12	19	31
-	8	2	10
Total	20	21	41

TABLE 3
ANALYSIS OF VARIANCE ON VOICE DOMINANCE SCORES

Source	df	Mean square	F	P
Voices	1	127.6	1.83	N.S.
Samples	1	155.1	2.23	N.S.
Voice x Samples	1	44.9	.65	N.S.
Within	37	69.6		
	40			

A point biserial correlation between relative scores and sex yielded $r_{pb} = .10$; $p > .10$. A test of the difference between the means of relative scores of males and females resulted in $t = 1.18$; 39 df ; $p > .10$. The relative scores employed in these tests had a possible range of -1 to $+1$ and were computed by subtracting the test score for the "punished" voice from that for the "rewarded" voice and dividing the difference by the sum of these scores.

No relationship was found between test scores and interview responses to questions regarding which voice was preferred or which voice was listened to most. The same was true of a similar analysis of *Ss* who both identified and preferred the same voice. Of the 41 *Ss*, seven discovered to some degree why money was being given, although none of the *Ss* discovered the purpose of the experiment to the extent of knowing which voice was connected with reward or that the test was a measure of the effect of training. Of the seven *Ss* mentioned, relative scores were correlated with an independent ranking of when *S* discovered why money was given, resulting in $\rho = .75$; $p = .05$. The discovery of the connection between voices and money was not significantly related to the over-all results ($\chi^2 = .04$; 1 *df*; $p > .90$). A total of 16 *Ss* believed that the experiment was an attempt to test their ability to concentrate. One *S* specifically said that *E* was trying to confuse him by distracting him from the real task of listening to the voices by moving the money.

D. DISCUSSION

The original prediction, that a significant number of *Ss* would give a greater number of verbal responses for the rewarded than for the punished voice in terms of content, is supported by the data. The results for 41 *Ss* ($p < .01$) indicate that there is a degree of relationship among the variables of reward-punishment and test scores. The result of a *t*-test of the difference scores of the test data shows that, with this more sensitive measure, there is an over-all significance ($p < .02$). Although the effect of rewarding voice "*D*" was not as pronounced as that of rewarding voice "*J*," results from both conditions are in the predicted direction (Table 2) and are not significantly different (Table 3).

Qualitative observations suggested that the money did, indeed, serve as "reward" for these *Ss*. Contrary to what happened in some pilot work with more sophisticated *Ss*, no *S* in the present experiment hesitated to take his pile of nickels. In fact, many pocketed the nickels after the training before *E* instructed them to do so. There were occasional remarks at the end of the experiment such as "I'll have cokes for a long time with these," or "I save nickels in my piggy bank." These kinds of reactions were usually followed by refusal to obtain larger denominations for their nickels although this was freely offered. The *Ss* occasionally remarked that they were "broke" and "could sure use the money." That *Ss* were interested in the money was further supported by the observation that *Ss* compared their winnings after the experiment.

The instructions and reward-punishment technique of the training were

planned to lessen the possibility that *S* would discover the design. Only seven of the 41 *Ss* reported in the interview that they had discovered why money was being given and taken away, but no *S* was able to indicate which voice had been rewarded or punished. No *S* discovered the purpose of the experiment. If *S* were aware of the connection between the reward-punishment and the voices during the training, it appears to have been a low level of awareness.

There was a nearly significant correlation between the test effect and the time during the training that *S* discovered why money was being given. The *Ss* who made the discovery early during the training gave more positive results than *Ss* who made the discovery later. However, the over-all results were not dependent upon mere discovery of this connection. Therefore, it is difficult to generalize about any relationship between awareness of the connection and the reward-punishment effect on the basis of this experiment.

The evidence suggests that the results are not predominantly due to the intrinsic attentional factors of the auditory stimuli. Although the effect of rewarding one voice was not as great as rewarding the other voice, the over-all effect is more than can be attributed only to the difference between voices. An analysis of responses to words spoken by one voice during a pause of the other voice indicates that these responses were infrequent and favored neither voice.

One may attempt to explain the results by supposing that the punishment had a greater distractive effect than the reward. If punishment had been used alone, this explanation would be quite tenable. However, in our training procedure the movement of coins for punishment was balanced with that for reward. The balance of distraction was furthered through the use of different modalities for the reward-punishment and ambiguous stimuli.

E. SUMMARY

As a test of the effects of monetary reward and punishment on responses to an ambiguous auditory pattern a total of 41 *Ss* were individually presented a training series of recorded sentences spoken alternately by two announcers. While one of these voices was presented, *S* was given a small amount of money and while the other voice was presented, money was taken away. This was followed by a test in which *S* was presented the two voices together in competition. The *S* reported at the end of each trial as much of the content as possible. As predicted, a significantly greater number of *Ss* produced larger scores for the previously "rewarded" voice than for the previously "punished" voice. A test of the difference scores for the responses to the re-

ward and punishment conditions was also significant. These results cannot be attributed to differences between the intrinsic factors of the auditory stimuli.

REFERENCES

1. JACKSON, D. N. A further examination of the rôle of autism in a visual figure-ground relationship. *J. of Psychol.*, 1954, **38**, 339-357.
2. ROCK, I., & FLECK, F. S. A re-examination of the effect of monetary reward and punishment in figure-ground perception. *J. Exper. Psychol.*, 1950, **40**, 766-776.
3. SCHAFER, R., & MURPHY, G. The rôle of autism in a visual figure-ground relationship. *J. Exper. Psychol.*, 1943, **32**, 335-343.
4. SMITH, D. E. P., & HOCHBERG, J. E. The effect of "punishment" (electric shock) on figure-ground perception. *J. of Psychol.*, 1954, **38**, 83-87.

Department of Research
The Menninger Foundation
Topeka, Kansas

Published as a separate and in *The Journal of Psychology*, 1956, **41**, 185-198.

LYSERGIC ACID DIETHYLAMIDE (LSD-25): XII. A PRELIMINARY STATEMENT OF ITS EFFECTS UPON INTERPERSONAL COMMUNICATION^{*1}

Departments of Medicine, Neurology, and Psychiatry of Mount Sinai Hospital, New York City; and the Bureau of Applied Social Research, Columbia University

H. LENNARD, M. E. JARVIK, AND H. A. ABRAMSON²

A. INTRODUCTION

Lysergic acid diethylamide is a drug which has been shown to produce a variety of changes in the intellectual (10, 13) and emotional functioning of non-psychotic adults (4, 11, 12, 14). Previous work in this laboratory has demonstrated that performance on objective tests measuring memory (8), concentration (5, 7) and motor abilities (1, 9) is impaired when subjects receive LSD-25. Some effects of this drug on social perception have been described by Hyde *et al.* (6).

The present paper attempts to describe the patterns of interpersonal communication of a small group of individuals under the influence of LSD-25. Our objective was to specify the dimensions along which interpersonal communication is modified and the dimensions along which it is unchanged when this agent is administered to several individuals in a group situation. The substance of this paper will consist of a discussion of a series of specific hypotheses which were formulated to explore the problem and is intended to serve as a preliminary report to be followed by a more intensive examination of this problem.

B. METHOD

1. Subjects

There were four paid, adult volunteers, three female and one male considered non-psychotic on the basis of a psychiatric interview and a battery of clinical psychological tests (2). Intelligence, as measured by the Wechsler-Bellevue Intelligence Scale, ranged from superior to very superior. Body weight ranged from 130 to 160 pounds. Age ranged from 22 to 26 years.

^{*}Received in the Editorial Office on October 15, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹This investigation was aided by a grant from the Geschickter Fund for Medical Research.

²We are indebted to Sandoz Pharmaceuticals, Inc., for supplies of LSD-25 and other compounds.

All subjects were college graduates, three in the field of education and one in business administration.

2. Procedure

a. *Experimental conditions.* At 9:30 A.M. two subjects received 50 micrograms of LSD-25 and two received 100 micrograms. The drug was given orally in 75 cc of tap water. Five months later a second experiment was conducted where the subjects were given a placebo consisting of 75 cc of tap water only. Since LSD-25 has no taste, odor, or color, subjects were unaware of the dosage they received.

At each session the subjects had eaten no food since the previous evening. One-half hour after the ingestion of the drug or placebo they ate a light breakfast. Neither smoking nor stimulants, such as coffee or tea, were allowed during the entire testing period. Prior to the drug experiment all subjects had participated in placebo experiments and two had received LSD-25. Between the drug and placebo experiments in this study there were several other LSD-25 sessions.

b. *Data obtained.* The following types of data were available for this preliminary analysis:

(1). Condition 1. A typed script from a tape recording of the above group, under the influence of LSD-25, discussing the topic, *The Place of Women in Society*.

(2). Condition 2. A typed script from a tape recording of this group discussing the same topic after they had received placebos.

(3). Condition 3. Typed scripts from tape recordings of group discussions in which members of our experimental group participated individually. The data used came from earlier sessions where the experimental member from the present group was under a placebo or "normal" condition while the others in his group were under the influence of LSD-25.

3. Analysis of Data

a. *Frame of reference for analysis of group communication.* A full exposition of our orientation with regard to group behavior must be postponed to a later date. At this time we shall briefly state our position. We agree with Bales that in studying group behavior, process, or functioning, what we are actually talking about is, "a sequence of qualitatively different activities of human beings, which is distributed in time and between individuals in a way that seems to be organized or patterned in a great number of ways" (3). Further, we believe that all interaction has "a problem solving direction

whether the people are trying to solve what we might call a substantive problem or not . . . the 'problems' involved are problems of interacting together and communicating" (3).

The study of the effect of LSD-25 on group behavior was made through the utilization of the most basic and easily quantifiable variables present. In order to arrive at generalizations of human behavior or the situation in which it takes place, we think it necessary to reduce stimulation and action into component parts which lend themselves to objective measurement.

b. Variables used. Two types of variables were used:

(1). Formal characteristics of communication: *quantity of speech* (number of words, number of lines, number of thought units); *direction of communication* (who initiates; amount of communication; to whom addressed); *other characteristics of communication* (interruptions, unfinished thought units, etc.).

(2). A system of categories devised by Bales (3) for the analysis of group interaction. Among the various methods of analysing interaction now available, the Bales system has been most widely used, thus providing us with a body of comparative data. The system of categories employed is described in Figure 1.

The set of categories used is "completely inclusive in the sense that every act which can be observed can be classified into one positively defined category" (3). The method is continuous in that every act is classified. The system further permits the scoring of "who to whom, in that for every act the author of the act (the actor) and 'target object' can be identified" (3). The target object is defined as "that area of focus of the situation (i.e., self, other, in-group or outer situation) which the actors aim to affect or change, or which is affecting or changing him, and to which he is, therefore, giving primary attention in the present momentary act" (3). We note that the unit of observation is the single act which is defined as "the single item of thought or the single item of behavior. Often the unit will be a single sentence expressing or conveying a simple thought" (3).

We should point out that our scoring deviated from Bales' method in that we classified responses from verbatim transcripts rather than from observation. The reliability of the scoring, computed for 100 acts, by inter-coding agreement, was better than 80 per cent.

In presenting our findings we shall first state the specific hypothesis explored, and the rationale for its formulation.

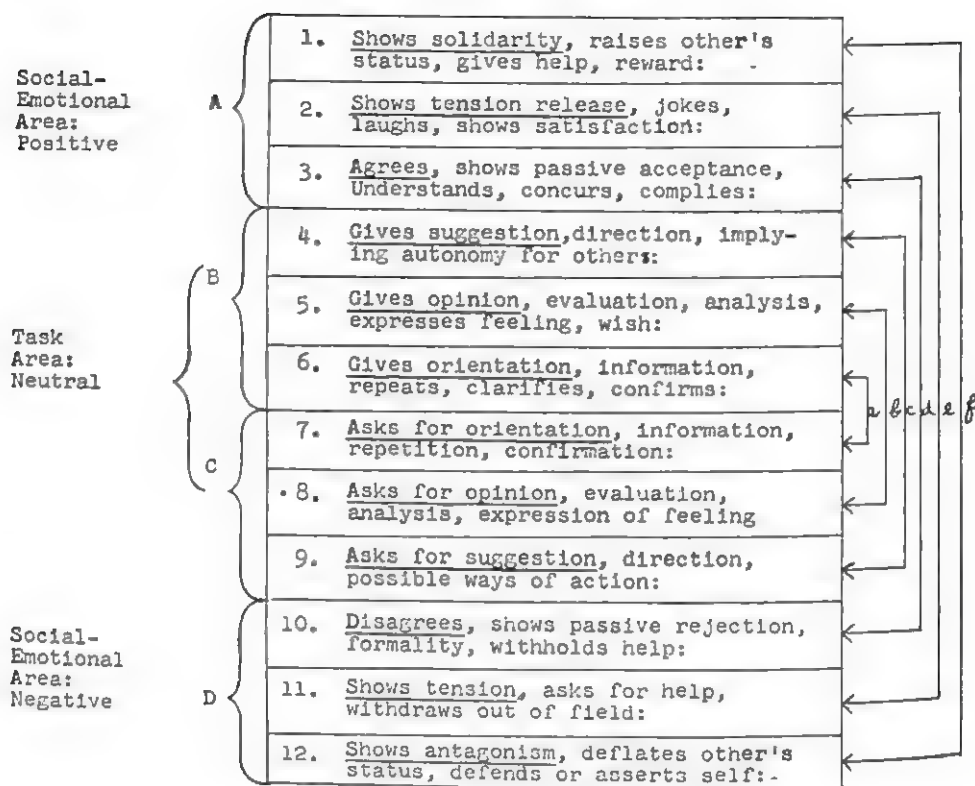


FIGURE 1
THE SYSTEM OF CATEGORIES USED IN OBSERVATION AND THEIR MAJOR RELATIONS
Key:

- a—Problems of Communication (6, 7)
- b—Problems of Evaluation (5, 8)
- c—Problems of Control (4, 9)
- d—Problems of Decision (3, 10)
- e—Problems of Tension Reduction (2, 11)
- f—Problems of Reintegration (1, 12)

- A—Positive Reactions
- B—Attempted Answers
- C—Questions
- D—Negative Reactions

C. RESULTS AND DISCUSSION

1. Hypothesis No. 1

The duration of proaction will be less under LSD-25 than under "normal" conditions; that is, the quantity of continuous speech for the group and/or the individual will be less under LSD-25 than under placebo or non-drug conditions.

Quantity of speech, one of the most basic variables of group communication, is easily quantified and throws light upon a basic group process. The above hypothesis represents our feeling that the impairment of intellectual

functioning, brought about by LSD-25 (10, 13), would be reflected in a shortening of individual verbal participation.

The measure of quantity of speech used was the number of typed lines. Although the number of words or number of propositions (syntactical idea units) could also have been used, the former is more easily obtained. Unpublished studies by one of the authors (H. L.) have established the correlation of lines and propositions to be approximately .9, the correlation of words and lines is probably higher still.

TABLE 1
AVERAGE DURATION OF PROACTION* FOR INDIVIDUALS AND GROUPS
UNDER THREE CONDITIONS

Condition	Individual				Average of group
	A	B	C	D	
1. Average proaction during LSD-25 in study group	.8	.9	.9	.8	.85
2. Average proaction during normal in study group	1.3	.9	2.0	1.2	1.4
3. Average proaction during normal while individuals participated in different groups where others had LSD-25	2.3	1.4	3.1	1.6	2.1**

*Proaction is defined as the ratio $\frac{\text{number of lines}}{\text{number of initiations}}$

**This is not really a group average. It represents the average proaction of the four individuals, derived from their individual participations in four different groups.

Table 1 gives the average duration of proaction under three conditions. Several observations can be made from the data:

1. Our hypothesis is borne out. Seven of the eight differences in proaction are in the expected direction; some of these differences are considerable.
2. The increase in proactive duration has to be explained in terms of at least two factors: the effect of LSD-25, and the effect of the group in which the individual plays a part.
3. No relation between LSD-25 dosage and decrease of proactive duration emerges from this data. Subject *A* and Subject *D* had 100 micrograms each under Condition 1 while Subject *B* and Subject *C* had 50 micrograms each. Subject *B*'s proactive value does not decrease from Condition 2 while Subject *C*'s proactive value does. One explanation which comes to mind would ascribe the relative relationships to something one might term "individual tolerance or sensitivity" rather than to absolute dosage. Previous studies in this laboratory have demonstrated considerable individual variation in the number of symptoms elicited by the same dose of LSD-25 (2).

4. We observe further that the nature or composition of the group of which the individual is a member makes a difference in his proactive rate. In each case the highest value appears when the individual interacts in groups where he is under condition "normal" while some or most of the others are under varying dosage of LSD-25.

This appears to be a very interesting finding. For interpretation, we must refer to the concept of equilibrium commonly used in the study of interaction. The communication system of a group is postulated to represent an equilibrium so that if one unit of the system fails to do its part, other units of the system make up for it. The presence of individuals whose proactions are decreased operates to increase the proactions of those who are not affected. This would be in line with the observation that in communicating with a rather taciturn individual it is necessary to speak for longer periods if one wishes to have the conversation continue.

5. We might also note that there is far less inter-individual variation under Condition 1 than there is under Conditions 2 or 3.

2. Hypothesis No. 2

The density of speech is expected to be less under LSD-25 than under "normal."

This again is a hypothesis which represents an extrapolation of the intellectual impairment findings (10, 13). One way in which communication processes may be affected, we reasoned, was in their conciseness. Density is defined as the number of propositions per line. If complete idea units are communicated in few lines we call speech "dense"; if it takes a great number of lines to communicate the same number of ideas, the speech is lacking in density.

Table 2 does not bear out our contention. Rather, there is a slight tendency in the reverse direction, that is, speech is somewhat more dense under LSD-25 than under "normal" conditions. One should note that the temporal element

TABLE 2
DENSITY OF SPEECH UNDER LSD-25 AND UNDER "NORMAL."

Condition	Number of Propositions		Ratio
	Number of Lines		
Normal	699		1.6
	440		
LSD-25	708		1.9
	374		

enters into considerations of density. It may be that individuals talk considerably slower (Hypothesis 2) under LSD-25 than under "normal," and that it takes considerably longer to communicate an idea or proposition under LSD-25 even though it is done within the same or fewer number of lines as under "normal."

3. Hypothesis No. 3

The number of unfinished statements will be greater under LSD-25 than under condition "normal."

Table 3 indicates that this hypothesis seems confirmed. There is a tendency for communication to "trail off" under LSD-25 and not be concluded mean-

TABLE 3
STATEMENTS UNFINISHED BY GROUP MEMBERS UNDER LSD-25 AND UNDER "NORMAL"

Condition	A	Individual B	C	D	Group total
Normal	1	1	3	2	7
LSD-25	16	24	8	36	84

Note: Definition of unfinished statement: any statement in which either subject or predicate was missing, and not implied.

ingfully. This was noted in varying degrees in all of the subjects. We are thus dealing here with a definite interference with the form of the communicative act.

4. Hypothesis No. 4

The number of interruptions will be greater under LSD-25 than under "normal."

The results given in Table 4 show that the converse of the hypothesis is true. The relationship between interruptions and unfinished statements is complementary, but in an inverse fashion. It seems that the same factor operating to make communications concise and complete, operates to maximize interruptions. One may hypothesize that individuals will be most likely to interrupt someone else when they are clearly conscious of, and have succinctly formulated the ideas which they wish to communicate. We believe that the

TABLE 4
INTERRUPTIONS UNDER LSD-25 AND UNDER "NORMAL"

Condition	Individual (Interrupted)				Group total
	A	B	C	D	
Normal	19	9	22	22	72
LSD-25	2	2	0	9	13

findings on decreased proaction, increased unfinished statements, and decreased interruptions should be viewed jointly. They present an integrated picture of changes in formal communication behavior which occur under the influence of LSD-25.

5. Hypothesis No. 5

The ratio of expressive acts to instrumental acts will be greater under LSD-25 than "normal," that is, the ratio of Bales' Categories 1 to 3, plus 10 to 12, to Categories 4 to 9 will be greater under LSD-25 than "normal."

This is the first in a series of hypotheses which utilizes the system of categories devised by Bales, and in terms of which our data has been scored.

As previously mentioned, Bales holds that the small group constitutes "a functioning social system." There are four types of problems which are being met by the system: (a) adaptation to conditions of the external situation, (b) instrumental control over parts of the situation in the performance of goal orientated tasks, (c) management and expression of sentiments and tensions of the members and, (d) preserving the social integration of members with each other as a social collectivity. Responses (acts) can then be classified in terms of the type of problem with which they are concerned. Responses concerned with Problem 3 plus 4 are called *expressive-integrative*. The following categories seen in Figure 1 are subsumed under this dimension: 1, 2, 3, 10, 11, and 12.

Acts dealing with Problems 1 and 2 are combined under the heading of *instrumental-adaptive*. The following categories comprise this area: 4, 5, 6, 7, 8, and 9.

The hypothesis was formulated on the assumption that the drug action on individual participants would function to disturb the relationship between the expressive and instrumental acts in group communication. In particular, we believed that there would be more acts devoted to the expression and management of tensions and sentiments. Thinking simply in terms of a dichotomy between emotional-personal concerns and task-centered content activity, it seemed that the scales would be tipped heavily in favor of the former, under Condition 1. This hypothesis was not confirmed as seen in Table 5. Yet, this very non-confirmation raises some very interesting problems. Can it be that the lack of change in the relationship between expressive and instrumental activity implies that the social functioning of the individual under LSD-25 is far less impaired than one may suspect from the evidence accumulated through individual testing?

For this group about the same quantity of behavior is devoted to tension

management as to problem exploration under both conditions. The distribution of specific categories undergoes marked changes but it is interesting to observe that ratios along the major analytical dimensions remain the same.

TABLE 5
THE RATIO OF EXPRESSIVE OVER INSTRUMENTAL ACTIVITY UNDER
"NORMAL" AND UNDER LSD-25

"NORMAL" AND UNDER LSD-25		
Condition	Number of Expressive Acts	Ratio
	Number of Instrumental Acts	
Normal	200	.40
	498	
LSD-25	208	.42
	500	

It is impossible here to explore fully the implication of a finding such as this. One would wish to gather further data to obtain additional confirmation of this finding, and to evaluate the relationship between emotional and task behavior within an analytical scheme other than Bales'. Nevertheless, here is an unexpected finding which deserves the attention of those presently doing pharmacological research and may well raise significant issues in the whole area of relating individual-psychological to social group processes.

6. Hypothesis No. 6

The ratio of Bales' Categories 1 to 12 will be less under LSD-25 than under "normal," i.e., the ratio of solidarity acts to antagonistic acts will be lower under LSD-25 than under "normal."

In devising this hypothesis we were operating on the assumption that the drug in its effect upon the individual would free repressed ideas, and would enable him to express antagonistic and hostile feelings more freely. Our premise was based upon previous observations of subjects under LSD-25 as well as upon reports in the literature (6). The actual relationship emerging is of a strikingly different order than envisaged. Table 6 shows that the ratio of solidarity acts to antagonistic acts is considerably higher under LSD-25 than "normal."

The table also shows the limitations in formulating hypotheses about a group process if one proceeds from the point of view of individual dynamics. Thus, based upon how an individual would be affected by LSD-25, we extrapolated to group behavior. We did not consider the type of group and the situational problems faced by the group. It is, however, in terms of these latter considerations that the finding becomes meaningful.

Let us consider the "group problem" of the four subjects involved. Four individuals, volunteering as subjects for experiments to test the effects of the new drug, meet in a strange room. They face the common threat of an

TABLE 6
RATIO OF SOLIDARITY TO ANTAGONISTIC ACTS UNDER LSD-25 AND "NORMAL"

Study group condition	Solidarity Acts	Ratio
	Antagonistic Acts	
Normal	$\frac{11}{39}$.3
LSD-25	$\frac{21}{9}$	2.3

unknown situation and the problem of trying to communicate and interact under conditions they expect to be difficult. Their actions are, therefore, ones of mutual encouragement and of helping each other. There is an absence of acts which would destroy this mutual support situation.

In looking at the data in Table 7, we note that the value of every negative emotional category is lower under LSD-25 than under "normal." Thus,

TABLE 7
DISTRIBUTION OF BALES' CATEGORIES FOR STUDY GROUP UNDER LSD-25 AND "NORMAL"
(Group discussion: *The Place of Women in Society*)

Category	Normal Percentage of total	LSD-25 Percentage of total
1. Shows Solidarity	1.7	3.0
2. Shows Tension Release	3.3	13.8
3. Agrees	8.2	7.9
4. Gives Suggestion	1.3	3.5
5. Gives Opinion	48.6	32.5
6. Gives Orientation	12.9	21.0
7. Asks for Orientation	2.9	4.7
8. Asks for Opinion	4.3	7.9
9. Asks for Suggestion	1.3	1.0
10. Disagrees	8.7	2.4
11. Shows Tension	1.3	1.0
12. Shows Antagonism	5.7	1.3

though the number of acts within individual categories are rather small, a total picture emerges which possesses some stability. Other collateral documentation of this hypothesis is available in the finding with regard to Hypothesis No. 7.

A word of caution should be introduced here. There is some evidence from Bales' work that negative emotional reactions increase when the same groups come together repeatedly. In our study, the LSD-25 experiment took place

prior to the placebo experiment, thus this factor may be operative here. This increase in negative reactions is explained by Bales partially as a function of the tensions generated by group decision, control, and status differentiation. One could account for it also by the common sense consideration that as people get to know each other better they feel less inhibited in expressing negative and antagonistic feelings.

7. Hypothesis No. 7

There will be more evaluation than orientation under LSD-25 than under "normal," i.e., the ratio of Categories 5 plus 8 to 6 plus 7 will be higher under LSD-25 than under "normal."

Categories 5 and 8 embrace those acts having to do with the giving of and asking for opinion, the giving of and asking for evaluation, and the giving of or asking for an expression of feeling or wish. All acts of this type, according to Bales, deal with the problem of evaluation. Categories 6 and 7, on the other hand, deal with the giving of and asking for orientation and information.

Here again, the originally stated hypothesis was derived from results obtained from the individual under the influence of the drug. We originally felt that the reactions of a subject would take on an increasingly personal coloration, be less objective, and revolve less around systematic communication.

After analysis of Hypothesis No. 6, however, we reconsidered the above formulation. Our involvement with extrapolation from individual processes had deceived us, and we felt that a reframing of this hypothesis in terms of a group orientation might be profitable. If the explanation with reference to Hypothesis No. 6 is valid then the primary problem faced by the individuals in the LSD-25 group was the shared uncertainty of having to function under the influence of LSD-25. Therefore, we would expect that acts concerned with the giving of and asking for orientation and information would increase. Just as the group had maximized behavior for emotional reassurance, we felt they would tend to maximize behavior designed to obtain cognitive reassurance.

It is the latter hypothesis developed around the group's problem which seems confirmed by the data, as shown in Table 8. The ratio for the LSD-25 group is less than half of that of the normal group. Looking at the absolute numbers involved, we observe that in the LSD-25 group, evaluative acts decrease and orienting acts increase.

The findings with reference to the last two hypotheses then may be considered complementary. They represent different aspects of the way in which the experimental group reacts to the administration of the drug, and to

the problem presented to it, i.e., (a) communication and interaction with one another under the drug, and (b) communication, particularly with reference to an assigned topic, *The Woman's Place in Society*. Here, as in the examina-

TABLE 8
THE RATIO OF EVALUATIVE OVER ORIENTING BEHAVIOR FOR THE
STUDY GROUP UNDER TWO CONDITIONS

Condition	Evaluation Orientation	Ratio
Normal	<u>370</u> 110	3.4
LSD-25	<u>286</u> 182	1.5

tion of previous hypotheses, we have come up against unexpected results, which, if they are confirmed in further work, strongly suggest the need for an intensive study of group process under LSD-25 to supplement the study of individual effects of the drug.

8. Hypothesis No. 8

The ratio of questions (Categories 7 to 9) to answers (Categories 4 to 6) will be higher under LSD-25 than under "normal."

The system of categories devised by Bales further permits characterization of acts, falling within the instrumental-adaptive dimension, as questions and answers. In task-centered communication the balance between questions asked and answers received is an important one. Every question initiates a chain of answers until the rise of additional questions.

We reasoned that a disturbance in the communication processes between individuals attributed to LSD-25 might result in unbalancing the equilibrium between questions and answers. Thus, there would be relatively more questions needed for a given number of answers under LSD-25 than under "normal."

This hypothesis seems confirmed, as shown in Table 9. In the communica-

TABLE 9
RATIO OF QUESTIONS TO ANSWERS UNDER LSD-25 AND "NORMAL"

Condition	Questions Answers	Ratio
Normal	<u>59</u> 439	.13
LSD-25	<u>96</u> 409	.23

tion process under LSD-25 there is more input in the form of questions needed for a given quantity of output (answers).

Let us briefly relate the finding presented here to those previously discussed in relation to Hypotheses 5 and 6. We noted in our discussion of Hypothesis No. 5 that the basic relationship between emotional and task-oriented behavior did not seem to be impaired under LSD-25. However, we observed in the findings with respect to Hypothesis No. 6 that the quality of the expressive-integrative behavior (or social-emotional area) did undergo an unexpected change. We now see a similar internal change within the instrumental-adaptive area with respect to questions and answers.

D. SUMMARY AND CONCLUSIONS

We wish to stress once more that the study reported was an exploratory one. Research along the lines described is continuing. The findings enumerated should, therefore, be considered more in the nature of working hypotheses for further exploration.

Some of the patterns of group communication under LSD-25 identified in our exploratory study were as follows:

1. Verbal output on the part of group members under the influence of the drug was restricted or shortened.
2. In groups where some members had been administered LSD-25 and others had not, there was a tendency for those who had not received the drug to increase their communication output. We interpreted this to mean that the individuals not under the influence of the drug made up for the deficiency in verbal communication of those who were.
3. When all group members were given LSD-25 there was a marked reduction in negative interpersonal responses. We suggest that disruptive social behavior is reduced spontaneously when group members operate under a common threat which, in our case, was the awareness of having to function socially while under the effect of LSD-25.
4. Despite the relatively high dosage of LSD-25 administered, the ratio between amounts of task activity and socio-emotional activity does not differ for the group when under condition "LSD-25," and when under condition "normal." The stability of this relationship under varying dosages and conditions deserves further examination. We suggest that the pattern of group communication seems less impaired by the drug, than changes observed in individual functioning had led us to expect.
5. The ratio of questions to answers, as well as to the ratio of orientation to evaluative responses, is higher in the group under condition "LSD-25"

than under condition "normal." We interpreted this as an attempt on the part of the group to restore cognitive clarity despite the felt impairment.

REFERENCES

1. ABRAMSON, H. A., JARVIK, M. E., & HIRSCH, M. W. Lysergic acid diethylamide (LSD-25): VII. Effect upon two measures of motor performance. *J. of Psych.*, 1955, **39**, 455-464.
2. ABRAMSON, H. A., JARVICK, M. E., KAUFMAN, M. R., KORNETSKY, C., LEVINE, A., & WAGNER, M. Lysergic acid diethylamide (LSD-25): I. Physiological and perceptual responses. *J. of Psych.*, 1955, **39**, 3-60.
3. BALES, R. F. *Interaction Process Analysis*. Cambridge: Addison Wesley Press, 1950.
4. CONDRAU, G. Klinische Erfahrungen an Geisteskranken, mit Lysergsaure-diethylamid. *Acta Psychiatr. & Neurol. Scand.*, 1949, **24**, 9-32.
5. FISHER, R., GEORGI, F., & WEBER, R. Psychophysische Korrelationen: VIII. Modell-versuche zum Schizophrenieproblem. Lysergsaurediethylamid und Mezcalin. *Schweiz. Med. Wschr.*, 1951, **81**, 817-837.
6. HYDE, R. W., VON MERING, O., & MORIMOTO, K. Hostility in the lysergic psychosis. *J. Nerv. & Ment. Dis.*, 1953, **118**, 266-267.
7. JARVIK, M. E., ABRAMSON, H. A., & HIRSCH, M. W. Lysergic acid diethylamide (LSD-25): IV. Effect on attention and concentration. *J. of Psych.*, 1955, **39**, 373-383.
8. ———. Lysergic acid diethylamide (LSD-25): VI. Effect upon recall and recognition of various stimuli. *J. of Psych.*, 1955, **39**, 443-454.
9. KATZENELBOGEN, S., & FANG, A. D. Narcosynthesis effects of Sodium Amytal, Methedrine, and LSD-25. *Dis. Nerv. Syst.*, 1953, **14**, 85-88.
10. LEVINE, A., ABRAMSON, H. A., KAUFMAN, M. R., & MARKHAM, S. Lysergic acid diethylamide (LSD-25): XVI. The effect on intellectual functioning as measured by the Wechsler-Bellevue intelligence scale. *J. of Psych.*, 1955, **40**, 385-395.
11. LEVINE, A., ABRAMSON, H. A., KAUFMAN, M. R., MARKHAM, S., & KORNETSKY, C. Lysergic acid diethylamide (LSD-25): XIV. Effect on personality as observed in psychological tests. *J. of Psych.*, 1955, **40**, 351-366.
12. MAYER-GROSS, W., MCADAM, W., & WALKER, J. W. Psychological and biological effects of lysergic acid diethylamide. *Nature*, 1951, **168**, 827.
13. RINKEL, M., DESHON, H. J., HYDE, R. W., & SOLOMON, H. C. Experimental schizophrenia-like symptoms. *Am. J. Psychiat.*, 1952, **108**, 572-578.
14. STOLL, W. A. D-lysergic acid diethylamide (LSD-25), a phantasticum of the ergot group. *Schweiz. Arch. f. Neurol. u. Psychiat.*, 1947, **60**, 1-45.

133 East 58th Street
New York City 22

LYSERGIC ACID DIETHYLAMIDE (LSD-25): XIX. AS AN ADJUNCT TO BRIEF PSYCHOTHERAPY, WITH SPECIAL REFERENCE TO EGO ENHANCEMENT*

The Biological Laboratory, Cold Spring Harbor, New York, N. Y.

HAROLD A. ABRAMSON

A. PURPOSE

It has been previously reported (1) that certain desirable reactions occur when LSD-25 is employed as an adjunct to psychotherapy. These included: (a) pharmacologic safety, (b) effectiveness in small doses, (c) conscious co-operation on the part of the patient, (d) elimination of the difficulties of operation on the part of the patient, (f) absence of ad-narcosynthesis, (e) feasibility of repeated administration, (g) excellent recall of events and ideas during psychotherapeutic interview.

In these earlier papers the preparation and management of the ambulatory patient has been discussed in detail. Emphasis was placed upon the fact that the therapist should anticipate that the integrative functions of the patient's psyche will aid the therapist in regulating the procedure of the interview. Other details have been previously reported and need not be repeated here. References to pertinent papers are appended (2, 5, 6, 7, 8, 9, 10, and 11).

It is the purpose of this communication to lay stress on the nature of the ego enhancement which occurs simultaneously with the ego depression so that during the LSD-25 reaction advantage may be taken of the integrative functions of the ego to provide insight. It is the function of the therapist to recognize the process of ego reinforcement and to employ it for the benefit of the therapeutic goal.

All of the subjects and patients discussed here are non-psychotic and ambulatory.

B. EGO DEPRESSION

It has been repeatedly reported that ego depression occurs when the administration of LSD-25 produces a psychotic state. Undoubtedly this does occur, but in a therapeutic interview with small doses ego depression is not as marked as with higher doses. The depressive process may be readily counterpoised with small doses of LSD-25 if the therapist manipulates the

*Received in the Editorial Office on October 15, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

interview situation to utilize the simultaneously occurring processes of ego reinforcement. With doses greater than 50 micrograms by mouth, the depressive processes may supervene. In the following verbatim report Subject S, who responds with very few symptoms when given 100-micrograms of LSD-25, shows marked ego dysfunction with withdrawal when this dose is increased. Following is a typical example of a verbatim recording of an experiment on Mr. S where ego depression is the main feature. However, the subject was being used in a psychological experiment and since no therapeutic intent was present during the LSD-25 reaction this may have influenced the nature of the ego response. This should be compared with the verbatim recordings elsewhere in this communication, of two other subjects who received smaller doses. In these two subjects, ego enhancement as part of the reaction is well illustrated.

Subject—Mr. S.
July 27, 1954

LSD-25 (orally)
200 micrograms

*(The questioning was carried out in the presence of two
observers, Mr. C and Dr. K.)*

Dr. When did you first notice your symptoms?

Mr. S. My first symptom was 15 minutes after ingestion of the drug when I was on Lexington Avenue and 59th Street.

Dr. You went out on an errand?

Mr. S. Yes, as soon as I took the drug.

Dr. This was the first time you had ever gone out on an errand while under the influence of the drug, wasn't it?

Mr. S. Yes.

Dr. You had previously been in a more protected environment.

Mr. S. Yes, that's true.

Dr. How did you feel when you were out at 59th Street and Lexington Avenue? How did you feel about the effects of the drug coming on?

Mr. S. Well, I was headed back here and I didn't feel anxious about the effects coming on.

Dr. When you were here, how did you feel?

Mr. S. Ah, well, I started feeling drowsy at the time.

Dr. Is that your usual reaction?

Mr. S. Usually, yes.

Dr. And you usually do feel drowsy.

Mr. S. That's true.

Dr. Let's see; it's four hours since you've had the drug now.

Mr. S. Right.

Dr. How do you feel now compared with your feeling three hours ago?

Mr. S. Oh, quite different now.

Dr. What is the difference?

Mr. S. Well, I feel more alert and able to think, and so on.

Dr. How did you feel at the height of your reaction?

Mr. S. Ah, everything seemed to be a void, shapeless forms, sort of amoebic in shape. And, ah, I felt as if I was, ah, suspended in their midst, neither moving, not being a bump on a log, nor a part of them, just suspended, and watching these shapes move around. There were no, no words had any significance of any kind.

Dr. Did you ever feel that way before?

Mr. S. Well, not usually.

Dr. You said that you—how did you feel about people? How did you feel about me?

Mr. S. There was no reaction at all. I felt completely inert.

Dr. Did you feel that way about Mrs. B. (secretary)?

Mr. S. Yes, the same way about her.

Dr. How about Dr. K.?

Mr. S. Same with him.

Dr. Would you say that you didn't give a damn about anybody, or that you couldn't give a damn?

Mr. S. I think it was that I couldn't give a damn because I didn't; the significance of the reality wasn't there to me at the time.

Dr. What do you mean, the the significance of the reality wasn't there? Would you mind describing that in more detail?

Mr. S. They were meaningless. I heard the words said, but I didn't know what they were. They were just meaningless.

Dr. You heard the words?

Mr. S. I heard the words, but they were just meaningless.

Dr. I asked you if you wanted to go down with me to the cafeteria about an hour ago, or an hour and a half ago, but you refused. Why did you refuse?

Mr. S. Well, I felt like staying up here and just lying down and resting.

Dr. So you had some feelings about things. You felt you didn't want to go down, so going away had some meaning to you, didn't it?

Mr. S. I don't know. It was quite confusing at the time, too.

Dr. In what way was it confusing?

Mr. S. Well, I couldn't orient my mind, for one thing, and mental confusion and, of course, the meaninglessness of things at the time.

Dr. What would you say was more meaningless, people or things?

Mr. S. I would say both. Everything was just shapeless.

Dr. Were you aware that you were doing a, were in the middle of an experiment?

Mr. S. I don't think I was at the time.

Dr. Oh, really? You mean you had really lost touch with what was going on?

Mr. S. Yes.

Dr. This happened to you once before, didn't it?

Mr. S. I don't recall now.

Dr. Remember the time you thought you were being washed down the toilet?

Mr. S. Oh, yes, but this wasn't that kind of reaction at all. There was nothing nightmarish at all about today's experiment.

Dr. Was this a pleasant feeling that you had, or was it unpleasant?

Mr. S. It was neither. It was just inert, "bump on the loggish" sort of feeling, with no emotion at all to it, vegetating, and so on.

Dr. Did you have any unusual fantasies?

Mr. S. No, this thing was an amoebic world, ah, the colors, well, there was no color either, just gray and black.

Dr. You mean you didn't see the colors in the room, the red books and the blue books?

Mr. S. No, I may have been asleep at the time.

Dr. Has your sense of color returned now?

Mr. S. Oh, yes, I know where I am now. I'm talking into a microphone.

Dr. And you had a good lunch?

Mr. S. Yes.

Dr. Or was it a good lunch?

Mr. S. It was good, yes.

Dr. Is that your dessert outside?

Mr. S. Yes, I'm going to have that.

Dr. How do you feel about Mr. C. being present at this time?

Mr. S. Well, I don't mind now, because I know who it is.

Dr. Did you mind him before?

Mr. S. Well, I probably wouldn't have paid any attention to him.

Dr. If you hadn't known who he was, would you have minded?

Mr. S. I don't think I would have had any emotion, one way or another.

Dr. How about now?

Mr. S. Well, now I wouldn't mind him being here. It's definite I wouldn't mind him being here.

Dr. Do you have any other questions?

Dr. K. I was just wondering; he said he may have fallen asleep. Did you say you were asleep at the time?

Mr. S. Well, I didn't know whether I was asleep. I don't think I was asleep at the time. I think Dr. Abramson was talking to me.

Dr. K. Would you say that you were slightly unconscious?

Mr. S. I would say slightly unconscious; we'll put it in that term.

Dr. K. And how often have you had LSD-25 before?

Mr. S. Oh, about 14 or 15 times.

Dr. K. Now, let's see, was it the same dosage each time, or did you increase the dosage?

Dr. Well, he's had a similar dose before.

(Six hours after ingestion of the drug)

(The subject listened to the foregoing recording)

Dr. It's now six hours since you've been given LSD. How do you feel about your reactions, now that you've listened to them on the recorder?

Mr. S. Well, I seem very, very far away now, and I'm glad that I recorded it before rather than now, when they aren't as vivid as they were at that time.

Dr. Why are you glad you recorded it then, rather than now?

Mr. S. Well, to record the results of how I felt at the time.

Dr. You mean, you would have been unable to describe the feelings now?

Mr. S. Well, I think, as clearly as I did then.

Dr. I don't understand.

Mr. S. I don't think I'd be able to describe it now as clearly as I did at that time.

Dr. Why not?

Mr. S. It seems to be quite different now, far away and now that I'm back here in the room—

Dr. You were here in the room when you described them.

Mr. S. But it was closer to the time when the reaction was taking place, at that time.

Dr. How far over the effects of the drug do you think you are now?

Mr. S. I think I'm about all over them now.

Dr. Really? All over them?

Mr. S. Except for some residual drowsiness.

Dr. Well, this is your typical drowsiness.

Mr. S. Yes, the effects of the drug.

Dr. The severe withdrawing sensation which you've had before?

Mr. S. Yes, I think so.

Dr. K. Was this a typical LSD effect for you?

Mr. S. No. I think it was different.

Note: This experiment differed from preceding experiments. In previous experiments Mr. S had always been in a protected environment during the onset of the LSD-25 response. Here his symptoms began rather rapidly. Fifteen minutes after he had ingested the drug he was alone on the street and noticed the onset of the LSD-25 response.

Study of the preceding verbatim interview discloses that the subject was essentially incapable of communicating effectively with the interviewer during the height of the LSD-25 response. This is an extreme example of ego depression under a fairly large dose of LSD-25. A similar withdrawal reaction has been observed in another male subject following 150 micrograms given orally. The short answers of Mr. S depend upon the prodding given to Mr. S to answer the interviewer. Study of his responses discloses that no one response contains more than 33 words. On interviewing Mr. S a year later he was able to recall spontaneously most of his general feelings and his experiences connected with the withdrawal process but he had to be reminded of the details from the verbatim recording.

C. EGO ENHANCEMENT

The following verbatim recording is from a well educated, psychoanalyzed subject who was emotionally upset because of her relationship with her divorced husband. The dose of LSD-25 was in the therapeutic range (20 micrograms). Mrs. X ordinarily would not speak as freely of her personal affairs. Under the drug, however, *stimulated by the questions of the interviewer*, she was able in the first part of the interview to kill off her former husband with a great deal of gaiety and subsequently in the interview she stated that he was of no concern to her. At the end of the interview, as well as three weeks later in the follow-up, Mrs. X volunteered that she had a feeling of complete relaxation never before experienced during the LSD-25 reaction. Compare this type of reaction with that of Mr. K which follows later in this communication. In Mr. K's interview, analysis of a dream led to ego reconstruction with new developments of a favorable nature in his personal and business life, whereas in the case of Mrs. X, the therapeutic process was more in the nature of catharsis.

Subject—Mrs. X.

LSD-25 (orally)

20 micrograms

Dr. Would you like to talk about Poppa (a nickname for the patient's divorced husband)?

Mrs. X. I don't know whether I would or not. I've talked about it and it's silly. I'm curious, because I'd like to know what's wrong with him. I'm sure there's something wrong with him. I'm sure he has some kind of mental disorder. I'd just like to know what it is.

Dr. Why would you like to know what it is?

Mrs. X. I'm just curious.

Dr. But what difference would it make?

Mrs. X. I don't know. I just want to know. I want to be able to name it.

Dr. Well, how do you feel about him aside from curiosity and wanting to give it a name?

Mrs. X. I hate him.

Dr. You hate him?

Mrs. X. I certainly do. I hate him with a venom (chuckles).

Dr. What kind of a venom?

Mrs. X. I can kill him off with pleasure.

Dr. Let's hear about that!

Mrs. X. I don't think it would be any great loss to anyone, even himself (coughs). He's a very unhappy person, and anybody that comes in contact with him is made unhappy, so—

Dr. So?

Mrs. X. He would be no great loss (laughs) to the world in general.

Dr. How do you know everybody feels about him the way you do?

Mrs. X. Well, I don't know about everybody, but I do know an awful lot of people who do.

Dr. Why do you hate him? Since we're conducting an experiment, I think we ought to find out why you want to kill him off. Or would you rather not talk about it?

Mrs. X. Oh, I don't care.

Dr. After all I might be very much interested in finding out what kind of venom you'd like to use.

Mrs. X. Well, I'd just like to kill him off. It would make my life much simpler.

Dr. Would it make your life simpler if you killed him off?

Mrs. X. I think so.

Dr. In what way?

Mrs. X. I don't seem to be able to get rid of him. He's always hopping into the picture in some unexpected way; never bringing happiness, never bringing anything pleasant, not to anybody in the family, particularly people of the family. Nobody in the family can ever have any contact with him without getting upset. He's a very upsetting influence, everywhere, anywhere, any time. And he is to me, too.

Dr. In what way is he upsetting to you, if you're willing to talk about it?

Mrs. X. He keeps fighting me through the children and through friends and acquaintances, and he puts the children in a very bad light with outsiders which I think is very unfair. And I can't understand it. I don't know what he's trying to gain by it. It's stupid. He's constantly losing; I simply cannot understand his point, the way he works. He confuses me and makes me very unhappy, because he's making life very difficult for the children. Whether he does it maliciously or not, I don't know. I've been able to give him the benefit of the doubt most of the time in the past, but now I question it. He talks a great deal about integrity and his code of ethics which is extremely strong according to him. And yet, he's done some of the most unethical things in the last few years, and I don't get his reasoning at all. He's very childish and creates a lot of trouble when there needn't be any at all. He gets everybody in an uproar for no reason.

Dr. Why is he able to do that? You've been divorced, now, how long?

Mrs. X. Well, actually five years, but I've been away from him seven or eight years.

Dr. How do you account for the fact that, although a long time has elapsed, you're still turned topsy turvy by him?

Mrs. X. Because he injects himself into our lives in a most unpleasant way, and it still upsets me emotionally when he does. There's such complete misunderstanding on his part. There's always a misinterpretation of motives and thought processes that I don't know how to

deal with him. Nobody knows how to deal with him. He's so unnatural and abnormal in everyday life reactions that you don't know how to take him.

Dr. What would you call abnormal in his reactions?

Mrs. X. Well, he hasn't a natural reaction to anything.

Dr. He hasn't a natural reaction to anything?

Mrs. X. Well, practically anything.

Dr. Well, would you give me an example of what you would call an unnatural reaction?

Mrs. X. Yes. People have congratulated him on being a grandfather. He has two reactions and they're pretty standard. One is, "I don't know anything about it." and, "They told me, I didn't know that I was a grandfather."

Dr. (Laughing) That's rather whimsical, isn't it?

Mrs. X. And at a great length he goes explaining to people that he didn't know he was a grandfather, that he didn't know where the children are living, that he didn't know anything about them at all, which of course is untrue.

Dr. You don't think he has his tongue in his cheek?

Mrs. X. Possibly, but I think he uses very poor judgment in pulling that kind of thing in the wrong place, at the wrong time, with the wrong people.

Dr. Well, maybe it's the right place, and the right time, and the right people for him.

Mrs. X. Possibly.

Dr. I don't see why that's unnatural. I can see that it's peculiar.

Mrs. X. Well, maybe that's what I meant. Then the other reaction is, "What's there to be excited about? It's just another kid. None of them are attractive till they're two or three years old. Maybe by then I'll be interested." Which I think is camouflage. He is really much more interested than he really will own up to.

Dr. But don't men very often take that point of view?

Mrs. X. Yes, I think so, but they don't carry it to the degree that he does. When it comes to the family, his children, the parents themselves, it's a little unusual, I think, to carry it to that extreme.

Dr. Couldn't one take his remarks as an example of understatement rather than an unnatural reaction? Couldn't that be his reaction?

Mrs. X. No, I don't think so. I think it's more than that. It's more than an understatement. There's always something added to it.

Dr. (Humorously) I don't think he deserves death at your hands for those feelings though, do you?

Mrs. X. Well, I only want to kill him because he only made me unhappy, that's all. And he makes my children very unhappy. Because of him they're pretty much distressed. They don't want to see him and they don't want anything to do with him. They have to, but they dread

it when they know he's coming. Just as I dreaded being around him the last few years.

Dr. You had the drug about an hour ago. How do you feel now?

Mrs. X. Well, like I did a little while ago; the edges of everything have been smoothed out.

Dr. What do you mean by that?

Mrs. X. Nothing seems quite as clear-cut as it was before; everything is a little, not exactly hazy, but instead of being squared off, things are rounded off.

Dr. For example?

Mrs. X. Angles don't seem quite, they seem just a little, to be a little less accurate.

Dr. Would you go around the room and describe it just as you see it now?

Mrs. X. Oh, everything looks perfectly natural. I don't mean that, it's just, like I feel often when I've had one drink too many, everything seems just a little hazier than usual. Otherwise, I see perfectly clearly.

Dr. Everything?

Mrs. X. I think so. Everything has a softer look. Colors aren't quite so strident. And the angles aren't quite as pointed. I just altogether have a softer feeling (chuckles).

Dr. Is it a good feeling?

Mrs. X. I don't know; I haven't decided yet. I'm not sure whether I like it.

Dr. Well, take something with an angle and describe it to me. Something that has lost its point and has become rounded as you put it.

Mrs. X. Well, they haven't actually become rounded. And I have to concentrate on an angle much more than I usually do to be sure that it's an accurate angle.

Dr. For example?

Mrs. X. Well, looking at the ceiling, the corner for example, I have to concentrate that they really are good angles.

Dr. Well, how do they look if you don't concentrate?

Mrs. X. Well, blurry, I mean everything just seems all white up there if I don't look very carefully. It's just the general feeling.

Dr. Well, would you mind describing the room without concentrating—just how things look if you don't concentrate?

Mrs. X. Well, the general feeling is one of warmth: the pink in the walls which means warmth; books that are very colorful; the lights. The whole room at the moment makes me feel warm. Otherwise, it looks very much as I've always seen it.

Dr. But you said that things are rounded.

Mrs. X. Well, not actually. My vision is still perfectly clear.

Dr. Would you like to walk and see that there is no great difference?

Mrs. X. I can assure you there isn't. That much I'm absolutely certain of. No. I would feel it if—

Dr. Would you like to walk some more?

Mrs. X. No.

Dr. Why not?

Mrs. X. I'm just quite content to sit. I feel a little lethargic. But that sometimes happens after I've had a drink.

(Five minute pause)

Dr. You didn't seem to walk very well.

Mrs. X. Well, it's an effort. I mean, I can walk a straight line, but it's an effort to do so. I feel tight, high, and I'm not interested in food at all. My lids feel very, very heavy. I'm liable to get very gay by the time this is over. I very often get that way when I'm high. I love to feel that way. But this feeling I don't particularly like. This is just the way I felt before I got seasick. My stomach isn't happy at all right now. I'm a little dizzy.

Dr. Could you walk, or would you rather not? I notice you have your eyes shut.

Mrs. X. My lids are very heavy. It seems I can hardly keep them open.

Dr. Well, do you want to sleep?

Mrs. X. I could, easily.

Dr. But yet you feel high, and you want to sleep at the same time. How do you account for that?

Mrs. X. I don't know. Of course I can keep my eyes shut, open, but it's with an effort. It's difficult to describe. I feel as I often do when I've had a little too much to drink. But I don't feel as exhilarated as I usually do. I feel a little more depressed in a way. I feel heavy and yet light-headed. I find this stimulated heartbeat quite unpleasant. I have the tendency to feel as though I were going to be short of breath. There's a slight numbness in my fingers. I think I can lift something up, but I think it would take an effort on my part to hold anything tight . . . This really has hit me. I feel very heavy and I don't like that feeling either.

Dr. Yet, you smile quite often when you say, "I don't like that feeling." So, it must not be all unpleasant.

Mrs. X. No, it isn't. It's pleasant in that I'm completely relaxed. I don't think there is anything about my entire body that is tense. I'm completely relaxed. But I can hardly keep my eyes open. It's very hard.

Dr. Well, if you shut your eyes, how do you feel? Does that make it easier to accept the effect of the drug?

Mrs. X. No. I think I feel it more with my eyes shut than with my eyes open. I generally feel very much gayer when I'm like this, and now I don't feel gay at all.

Dr. Yet, you were smiling when you said that.

Mrs. X. I know, but I'm not really gay. I don't feel sick or anything of the sort. I just don't care.

Dr. You just don't care about what?

Mrs. X. Anything.

Dr. Well, how do you feel about Poppa (divorced husband) now?

Mrs. X. He doesn't really disturb me. There's only one thing I want to say about him—to hell with him (laughs)! He really isn't worth worrying about.

Dr. Well, that isn't the way you felt about him 20 minutes ago.

Mrs. X. I know.

Dr. Well, how do you feel about him now?

Mrs. X. That's just the way I feel—to hell with him!

Dr. What do you mean to hell with him?

Mrs. X. Right now he couldn't worry me very much. He's really too far removed. I allow myself to get disturbed on the surface occasionally, foolishly, but actually he can't harm me any and he doesn't mean anything to me. I get annoyed with myself when I do get upset.

(Five minute pause)

Dr. Do you feel all right?

Mrs. X. You're really bothering me you know. If you would only leave me alone, I could be very comfortable and happy.

Dr. Why do you want to be left alone?

Mrs. X. I just want to relax; I feel perfectly comfortable. I don't particularly want to talk; I don't want to think. I'm completely relaxed, and I just don't want to be bothered.

Dr. You feel better lying down than you did sitting up, is that right?

Mrs. X. Yes.

Dr. What's the difference do you think?

Mrs. X. I think I'm more relaxed this way. . . . Do you know it's an effort for me to talk? It really is. I just want to be left alone (laughs). I don't want to be bothered talking or thinking.

Dr. Is it because there are many things that you'd like to say and you don't want to say? Do you think that has something to do with it?

Mrs. X. No. Definitely not, because why would I hesitate to say them here?

Dr. Well, because everything is being recorded.

Mrs. X. Oh, that wouldn't make any difference. I don't object to that at all.

Dr. Well, how do you feel about Poppa now?

Mrs. X. The same way I did a few minutes ago. He isn't worth worrying about.

(Short pause)

Dr. You've had the drug about an hour and 40 minutes ago. How do you feel?

Mrs. X. Very relaxed. Almost numb you might say, except that my eyelids still feel very heavy. It's very difficult for me to open my eyes and look around.

Dr. Would you like to stand up? I'll turn on the light.

Mrs. X. I don't like the light on.

Dr. Why?

Mrs. X. It hurts my eyes.

Dr. In what way?

Mrs. X. I have the tendency to open my eyes and I don't want to. They're much more comfortable shut.

Dr. It took you a long time to answer. Ordinarily you answer immediately.

Mrs. X. I feel settled down. I'm very relaxed but definitely slowed down. My mental processes feel that way too. It's much more of an effort for me to talk and to think and to express my feelings in detail.

Dr. Are you aware of everything going on around you?

Mrs. X. Perfectly.

Dr. What's your attitude toward taking the drug?

Mrs. X. I think it's an interesting experiment.

Dr. Why?

Mrs. X. I didn't think it would be able to produce an effect, this effect in a person. And I'm aware of everything that's going on.

Dr. Would you like to stand up now?

Mrs. X. I don't seem to be as sure-footed as usual.

Dr. Can you read all right?

Mrs. X. Oh yes, everything is perfectly clear. Except that I feel heavy-headed. I feel as though there were a load on my chest and I want to lie down again.

(Short pause)

Dr. . . . It's two hours now. How do you feel?

Mrs. X. I'm beginning to feel much more normal.

Dr. What's the difference?

Mrs. X. Well, I'm more tense than I was, and things are becoming sharper in my vision. It's less of an effort to think and much less of an effort to talk, and I'm slowly coming back to normal, and I don't know if I like it. I'm not nearly as relaxed as I was. I can feel myself getting tense. There's a kind of sense of disappointment in coming back. I'm coming back to reality so to speak. I felt a little unreal before. I felt floaty, as though I had taken a kind of a trip to the moon, and now I'm back again. It's a little disappointing to see everything as it was.

Dr. Why are you disappointed in reality?

Mrs. X. (Laughs) Well, it's a lot easier to feel the way I did before. I didn't care so very much about many things. I'm much more intense this way than I am the other way. It's a very comfortable feeling not to feel so intense.

Dr. But at one time you said you'd like to kill Poppa. At another time you said that you didn't care about him. Which was the intense feeling?

Mrs. X. The killing, of course.

Dr. Maybe the 'not caring' was just as intense.

Mrs. X. Maybe.

Dr. Do you think the drug's effect is wearing off?

Mrs. X. Yes, I think so. I don't think entirely, but some.

Dr. I see you are responding more quickly to my questions.

Mrs. X. I feel much more normal.

(During the third hour after ingestion of the drug)

Dr. How do you feel about Poppa now?

Mrs. X. I feel about him now the way I usually do.

Dr. How is that?

Mrs. X. I still hate him (laughs). And I still say it wouldn't be any great loss to anybody else or to himself if he were killed off.

Dr. It's two hours and ten minutes that you have been under the influence of the drug. Does it seem that long?

Mrs. X. No. It seemed like a much shorter time. And I don't feel the effect of it nearly as keenly as I did a while back. I don't recall ever having felt as completely relaxed as I did at a certain period after taking the drug, physically and psychologically.

Dr. Even after alcohol?

Mrs. X. Yes, even after alcohol. This was more complete. It was everything from my toes to the top of my head, physically and mentally too.

Dr. Why do you think you were mentally relaxed?

Mrs. X. I don't know, unless it goes along with the physically relaxed. I don't know. It must be the effect of the drug. I found it a little hard to concentrate.

(The subject returned to the office three weeks later)

Dr. It's about three weeks since you've had LSD-25. Looking back what would you say the most important reactions were?

Mrs. X. A feeling of complete relaxation which I have never felt before.

Dr. Even when you've had three or four cocktails?

Mrs. X. This was an entirely different kind of feeling. I've never experienced it before.

Dr. Was it pleasant?

Mrs. X. Very.

Dr. Would you like to re-experience it?

Mrs. X. Yes, I would.

Dr. Would you mind describing again what the type of relaxation was? It's perfectly all right to repeat what you said when you had the drug. I don't think any restrictions should be placed on your words.

Mrs. X. I wish I could describe it more minutely. I don't think I have anything further to say, except that it was so complete, that I was a little surprised by it. Never having felt that way before, I don't think I realized how complete it was until the tenseness began to come back again. It was a physical reaction. The tenseness began to come back in my fingertips and my toes and gradually through my body.

Dr. It was a pleasant experience would you say?

Mrs. X. Well, it wasn't unpleasant enough to have made it an unpleasant experience. I'd be quite willing to take it again.

The reports in the literature on the LSD-25 reaction have in general emphasized the disturbing effect of the drug on the subject's adaptive functions. According to most reports the ego structure which to all intents and purposes had been functioning suitably, after the ingestion of 50 micrograms of LSD-25, is disturbed by the more primitive parts of the subject's adaptive processes.

In the following experiment, Mr. K, who was a 35-year-old scientist working in a local hospital, and was charged with the construction of scientific equipment, took 50 micrograms of LSD-25 by mouth. He was the subject of an experiment lasting most of the day. The subject fell asleep after lunch and upon awakening described a dream he had about a greenhouse composed of spider webs. This dream led to recall of a repetitive dream of spider webs, which he had been having since he was 15 years old. This in turn led to the development of an important situation in connection with K's work problems. It appeared that at the hospital where he was employed he was unable to communicate easily with his superiors. A project would be assigned to him by his supervisor and without discussing the matter further he would vanish for a few months and then suddenly appear with the completed project. This led to difficulties in his work relationships. These difficulties carried down the line to his subordinates who also felt the uncertainty of his position.

The reader should bear in mind that the interview occurred in 1952, when I was not as certain as I am now of the safety of dealing with K's reaction to LSD-25 and the accompanying dream. If the dream had been reported by a patient in analysis without the LSD-25 reaction the procedure would have been somewhat different. Primarily, here the safety of K was the most important problem. Although K had had LSD-25 several times previously, I still wished to proceed extremely cautiously in exposing unconscious material. However, I do not believe that my anxiety about the interview was too great because I had been using LSD-25 for more than a year.

Subject—Mr. K
August 1, 1952

LSD-25 (orally)
50 micrograms

(Time of interview was two hours after the ingestion of LSD-25)

Mr. K. I've just been sitting here sound asleep after lunch. I put my head on the table and dozed off into a sound sleep until I was awakened by the sound of the telephone buzzer or the doorbell. After I heard the buzzer, I woke up enough to know that I had been dreaming and realized that it was important for me to remember what that dream was. It's going (the dream) rapidly, so here's what it was.

I FELT THAT I WAS TAKING PART IN SOMETHING IMPORTANT. I WAS IN A

TROPICAL CLIMATE AND I WAS OUT IN THE FIELDS OR IN THE JUNGLE WHERE THERE WAS A LARGE CLEARING, VERY BRIGHT OVERHEAD. AND ON THIS CLEARING THERE WAS SORT OF A GREENHOUSE COMPOSED OF SPIDER WEBS AND IT LOOKED LIKE A TENT. AND IT HAD A WOODEN RAILING AROUND IT, VERY LARGE. IT EXTENDED ABOUT 30 OR 40 FEET ON THE SIDE, I MEAN AROUND IT, MADE OF VERY WEATHER-BEATEN OLD BOARDS. AND IT WAS A TENT-LIKE STRUCTURE. AND THERE WERE TWO OTHER PEOPLE WITH ME; I DON'T KNOW WHO THEY WERE. WE STARTED TO WALK INTO THIS TENT AND THE LIGHT, FILMY MATERIAL LOOKED LIKE MOSQUITO NETTING. WE STARTED TO WALK INTO IT. IT WAS THICK AND TENUOUS AND IT HAD THREE DIMENSIONS. AND THE WHOLE INSIDE OF THE TENT WAS MADE UP OF LAYER UPON LAYER OF MATERIAL. WE STARTED TO WALK INTO IT, WAVING OUR ARMS, SLICING THROUGH IT, BREAKING THE SHEETS OVER OUR HEADS SO WE COULD GET INTO IT. AND AS WE GOT INTO IT I SORT OF HAD THE FEELING THAT WE HAD SOMETHING ABOVE US. WE CUT THROUGH LAYERS AND LAYERS OF THIS MATERIAL. IT WAS ABOUT 20 OR 30 FEET HIGH AND WE MOVED TOWARD ONE OF THE SLOPING SIDES. WE WERE APPROACHING NEARER THE TOP SURFACE OF THE TENT. WE GOT INTO IT WITH OUR ARMS. AND THERE WERE SPIDERS THERE. THEY LOOKED LIKE GRASSHOPPERS THOUGH, JUST GETTING CLOSER, SAY WITHIN FIVE OR SIX FEET OF THESE SPIDERS; THEY MUST HAVE BEEN ABOUT A FOOT HIGH. AND THEN THE DOORBELL RANG AND I STARTED TO WAKE UP. I STAYED IN THIS SORT OF STAGE AND THE DREAM WENT FURTHER. WE CUT MORE AND MORE AND WE GOT CLOSER AND CLOSER TO THE SPIDERS, UNTIL FINALLY THERE WERE JUST ABOUT TWO MEMBRANES SEPARATING ME FROM THEM. Then the secretary walked into the room and I woke up entirely.

Dr. Were you frightened in this dream?

Mr. K. I wasn't frightened.

Dr. You called it a nightmare when I asked you if you had been sleeping. You said you had a nightmare.

Mr. K. Yes, I would call this a nightmare. It was not a pleasant dream. Cutting through the material wasn't good. I tried to stay half asleep when I was telling the story. You asked if I was frightened. No, I wasn't frightened but it was unpleasant, coming so close to those spiders. It was just unpleasant, the tenuousness of and the feel of the material we were cutting through. Already the dream is fading. I'm almost forgetting it now. I just remembered I wasn't frightened; it was just unpleasant.

(The subject became quiet and had to be questioned actively.)

Dr. You were cutting something, the spider webs or the material, with nothing in your hands?

Mr. K. No, nothing in your hands, just sort of tearing it apart with your own hands, using the arms and hands rigidly extended like they were sticks of wood.

Dr. Are you tired now?

Mr. K. No.

Dr. Do you feel worried or concerned? You seem to be sighing as if you were pretty exhausted.

Mr. K. This sighing I think doesn't necessarily connote exhaustion. It's something that you do. Exactly this minute, I feel very relieved. I feel like I've rested. My hands and legs are not shaking as they were before. Since I'm not fully awakened I feel rather refreshed.

Dr. You said there was something bright overhead.

Mr. K. The bright overhead I think was the tropical sun. I saw it when I was in Central America.

Dr. You want to look up to the ceiling?

Mr. K. Well, there's the bright, not too bright, yellow glow overhead (electric light fixture). Here's another thing I didn't say before. I was standing underneath green trees, looking out into a bright, sunlit field which was covered with this tent.

(The subject continues to refrain from speaking spontaneously.)

Dr. Have you ever been in the tropics?

Mr. K. Yes.

Dr. Did you enjoy it?

Mr. K. Yes, I think I did. I always said that I enjoyed it and I think of it as being enjoyable.

Dr. Did you see any spiders?

Mr. K. No.

Dr. These weren't really spiders were they?

Mr. K. These were big, green spiders, like grasshoppers. I don't know why these didn't look like spiders, yet I called them spiders. I felt as though they were spiders and I assigned this name to them. It seems to fit.

Dr. What gave you the spider feeling?

Mr. K. The spiders had built this tent, this tent shaped structure; the whole thing was filled with this material. Although there were thick spider webs, they were just resting at the top surface of the tent.

Dr. Were they crawling?

Mr. K. No. I wish I could quickly sketch what this tent was like. I haven't described it too well.

Dr. Would you like to draw it?

Mr. K. I can't draw it. I wish I could. I wish I could paint it. I can't draw it though. I'd be too concerned with the technicalities of drawing. I wouldn't get the impression down. The material was woven thick like cotton and they were at the top of it.

Dr. Well, the idea was to get to them?

Mr. K. Well, it started out, I didn't know they were there. I started clearing out the tent and as I got closer to it, I realized there was something up there. I remember I just had to cut three or four membranes to get to them. I remember them as being like cotton wool stretched over the whole tent. The whole thing was this material we were cutting through. And I deliberately cut closer to them so I could see what they were.

Dr. Have you ever lived in a tent?

Mr. K. Yes, I've gone camping in tents.

Dr. Were you afraid to sleep in a tent?

Mr. K. No.

Dr. In the dream there were some other people with you.

Mr. K. The people with me, I think there was one on my left and one on my right.

Dr. You don't know who they were?

Mr. K. I don't know who they were. I didn't see their faces. I don't know what they were, men or women. I feel as though it was an exciting dream, a lot of action in it.

Dr. What does the dream bring to your mind?

Mr. K. There was one person on my right and one on my left. It was either the feeling that I knew I was waking up and I knew it was important for me to remember what the dream was. I didn't realize at the moment where I was. I did and I didn't.
(Since so few data are available on sleep during the LSD-25 reaction, it seemed desirable to obtain details of the waking process.)

Dr. Did you suddenly wake up when you heard the buzzer?

Mr. K. No. I heard the buzzer ring twice, I think, and I woke up slowly because I immediately tried to put myself back into this dream. I went back into it so that I would remember it.

Dr. Did you ever have a similar dream?

Mr. K. No. If I have similar dreams, and I've had them, they're identical dreams. But I've never had one like this.

Dr. And you don't feel disturbed and you don't feel relieved?

Mr. K. No.

Dr. How do you feel now?

Mr. K. I don't feel anything specifically. I still feel I'm willing to do anything and I'm in a very relaxed position as to my head and shoulders.

Dr. Are you very warm?

Mr. K. No, not particularly.

Dr. Were you warm in your dream?

Mr. K. Not consciously, not uncomfortably at least.

Dr. You're not warm, yet you seem to be gasping for breath.

Mr. K. I'm not warm, yet I am gasping for breath. This seems to be a reaction for which I'm looking.

Dr. Do you sometimes do that?

Mr. K. No.

Dr. Not even when you're nervous?

Mr. K. Not unless it's due to some physical exertion.

(The anxiety of the subject appeared to decrease. It appeared safe to proceed with free association and to attempt the interpretation of the dream.)

Dr. Now, anything that comes into your mind.

Mr. K. This I always find difficult. So many things come into your mind and you have to pick out one. The word "tent" reminds me of when I went camping down at Lake ——— last October with my wife

and two friends and another married couple that we know. This tent in my dream was similar to that except very, very much larger. The tents we slept in there had regular wood platforms and for sleeping there were cots. The tent in the dream was much larger although the same shape. Spiders bring to mind the identical dream I said I have had, not in the last few years, but possibly five times I've had this identical dream.

I AM IN THE UPPER CORNER VERY, VERY HIGH UP IN A VERY, VERY LARGE WEB. THIS IS A TWO DIMENSIONAL WEB, AND I AM AT THE APEX OF IT. ALL THE STRANDS ARE COMING TOGETHER TOWARD ME AND I'M ENTIRELY ALONE. THE WHOLE THING IS LIKE IN THE TOP CORNER OF A LARGE, GREAT BOX. I CAN SEE NOTHING OUTSIDE OF ME AND THIS WEB IS STRETCHING OUT FROM ME IN SLOPES, OUTWARD AND DOWNWARD IN 45 DEGREE ANGLES, AND IT GOES THOUSANDS OF FEET AWAY. IT IS A VERY, VERY LARGE NET AND WAY DOWN IN ONE CORNER—I CAN HARDLY SEE THEM—THERE'S A TERRIFIC TURMOIL OF PEOPLE, VERY, VERY CLOSE TOGETHER AND ALL STRIVING AND CLAWING AND MILLING AROUND TOGETHER, JUST LIKE YOU SEE ANTS IN A NET WHEN YOU TURN IT AROUND. AND ALL OF THEM DOWN THERE, THOUSANDS OF THEM, THEY'RE PILED HIGH. THEY SEEM TO BE ON TOP OF ONE ANOTHER, THERE'S NO FLOOR THAT THEY'RE ON; JUST A BIG BUNCH OF PEOPLE DOWN THERE, ALL PUSHING AND CRAWLING ALL AROUND EACH OTHER.

This is an identical (repetitive) dream which I've had. The earliest I remember it is somewhere about when I was 15. And this dream has been recurring probably once every three or four or five years. The jungle sky, or the Central American placement of the dream I possibly attribute to having lived in Guatemala for three years during the war. My stay there I think was very rewarding and very satisfactory. I enjoyed it there. I would not hesitate to return. The hot sun, the tropic sun is something you don't experience in a more northern climate. It's a close and warm feeling. The sunlight is brilliant, white, couldn't possibly confuse it with the yellow globe overhead. I don't have any particular feeling about the weather-beaten board railing around these which was about a foot off the ground. The impression is very clear in my mind; I could almost see the knots of the wood. I can't draw anything from the spiders and grasshoppers. Maybe we have a few spiders around the house and we have to kill them, step on them, sweep them out. I have a revulsion toward spiders, no more than usual.

Dr. The spiders weren't spiders?

Mr. K. No. I said I gave them that name; they felt like spiders. Perhaps it was the association with the spider web. I still think the name fits. A grasshopper is a descriptive term; a spider is not necessarily an evil thing or a horrible thing. It's an insect, something of the insect kingdom. They looked like grasshoppers. I could only definitely see the upper halves of their bodies; typical large grasshopper head, the short foreleg tucked under them, something like a praying mantis has—short forelegs. I would say it was undesirable to touch them. The end of the dream was at about 3:20.

Dr. There's one inconsistency in what you've told me. The time you were in Guatemala, you had your wife with you?

Mr. K. Yes.

Dr. You said your life in Guatemala was enjoyable, but from other things you've told me, you wanted a divorce from your wife during that period?

Mr. K. Yes. We had had a disagreement before we left this country for Guatemala. One of the things that I wanted to do before I left the country was to learn Spanish and the Company had arranged for a teacher to teach us. I was interested in taking it as much as time would allow, but my wife was not interested at all and she stated flatly she did not want to learn the language. She just didn't want to be inconvenienced to that extent. This was the time when we were disagreeing on almost everything. I had felt, although not actually, I had felt separated from my wife at that time and had been traveling in Maryland quite a bit for the Company and only returning home on weekends. During this period immediately preceding our departure about six months, we did feel very compatible. During the time we were in Guatemala she lost, I think it was, two of her pregnancies. She's had quite a bit of trouble holding a pregnancy in the past few months. I remember definitely her losing one there, and I think there must have been two. In spite of this I think that during my whole three years in Guatemala, generally speaking, my wife and I were staying together more for social and economic reasons than just because it was too difficult to break up.

Dr. 'Cotton wool,' what does that bring to mind?

Mr. K. I used the word 'cotton wool' to describe the layer upon layer characteristic of this tent. I've been using cotton wool recently in sterilizing areas of skin before giving a hypodermic injection to my wife so she can maintain the pregnancy she now has. Cotton wool is a very common tool in our house. I scrubbed my hands off with cotton wool the other day after a cat had bitten me. I got involved in a cat and dog fight. I had to separate the two animals and the cat bit me, and I cleaned it off with the cotton wool.

Dr. Have you and your wife had cat and dog fights?

Mr. K. Unfortunately, not too much cat and dog. Generally they have been more bitter and more subdued than an out and out scrap. We have never slapped each other's face—nothing ever physically. Very seldom do we even come to harsh words—spiteful, mean words. In fact, I don't think we've ever called each other names. For instance, we've had subdued, long standing arguments—feelings, emotions prolonged for days on end.

Dr. How about your relationship with your mother?

Mr. K. With my mother? That's hard to describe.

Dr. When you first had these dreams; when you were in the web.

Mr. K. Oh, I remember I was just a young boy, very young, about ten or fifteen, around there. My relationship with my folks has always been

pretty good. They've never resorted to harsh discipline and they've never beaten me.

Dr. But what was your relationship with your mother?

Mr. K. I always felt that she exerted a very strong moral influence on me, in that *children must be seen and not heard*. I must always do right, go to church, shouldn't lie or steal. I looked upon her as being a good woman, not utterly right in everything she says.

Dr. But if you were seen and not heard, you were restricted.

Mr. K. I was restricted but restriction is only teaching your children manners.

Dr. I think it's restriction. I have a feeling that in the dream you're going back in part to the characteristic repetitive dream which you had in childhood. It goes back to your relationship with your parents, probably your mother, to the time you were brought up to be seen and not heard. Maybe that accounts for the fact that you work as you do, quietly, you're seen and not heard; you don't like to talk to your superior about what you do. That was your pattern of adaptation which you've carried over into your life's work. And your relationship with your wife, in all probability, is influenced by your attitude toward your mother which you've never been able to face—your love for her. She restricted you. In a way you were in a spider web.

Mr. K. I was in the web; I was very remotely away in that corner, very alone. Utterly alone.

Dr. And I think this dream reflects your present difficulty at work. And as a matter of fact, I wouldn't be at all surprised if your anger at being dominated to a certain extent by your mother's attitude to be seen and not heard is carried over in your relationships with your wife, your mother-in-law and your supervisor, because that's exactly what your supervisor is complaining of: you're seen and not heard. Do you see?

Mr. K. Yes. Not that I'm not working, but that I'm not working the way he'd like me to.

Dr. You're alone in a web of your own. And the spider not being the spider—

Mr. K. Is my mother.

Dr. No. I wouldn't say is your mother. I don't have enough evidence for that, but if all—

Mr. K. It's the frustrations that I feel, the things, the troubles that I run into.

Dr. And everybody is making a web around you. Even I weave a web about you, and you feel—you remember that song, "Don't Fence Me In"?

Mr. K. Yes.

Dr. You're fenced in by your marriage, by your mother-in-law and—

Mr. K. I don't like to be restricted.

Dr. That's it. And that's what this dream represents. You feel a lot better now, I think, because you can see that that's one of the things

you've never been able to face; that you dislike the things you also want. You've been brought up to be fenced in. If you're not fenced in, you'll get panicky.

Mr. K. What do you mean, if I'm not fenced in I'm panicky?

Dr. Well, if a child is brought up to adapt himself to life in a certain way and if he gives up that way of life, he has a new type of adaptation to develop, and if you're alone and without your previous technique of adaptation, you'll worry, "How will I do?" For example, I think your relationship with your supervisor comes in. You're not so much concerned with discussing things with other people, but it's how your discussions will go, "Will you conduct yourself in a civil manner?", or will you, as you told me before, know when the fences may be removed?

Mr. K. There's a clue in what you've been saying. The thought is perhaps best represented in that if you took away my fences I'd become panicky. I know myself that I want to get rid of my fences; I wish I had the means. This thought occurred to me earlier today. I wish I were powerful enough sometimes so that I could overthrow authority around me, that I could tell people, "Beat it, I don't want to see you; you bother me; go away." I don't have that authority. I don't have that power. There are many things I wish I could do.

Dr. Many people would like to do that.

Mr. K. I'm certain of that. I'm glad there are not more of us. We can't all go and do it. I feel perhaps that I'm also trying too hard to be a success now, although I don't want to stop doing anything that I'm doing. There is something else I noticed. Immediately after this dream had been recorded, all my muscular tensions stopped. My hands and legs stopped shaking. Then I listened to the recording of it. Now I don't think I'm living through it as much, as time elapses, because after that I began to feel these symptoms again—the muscular tiredness. Before I went to sleep, while you were in the other room dictating some letters, I walked up and down through the corridor, and I know my emotions were very mixed but I can't describe them now.

Dr. Your emotions were mixed up about me, and about being a guinea pig and that you were in this web of psychological—?

Mr. K. No, but there was a little—

Dr. Do you think I was the beetle who was going to take you apart?

Mr. K. No. I don't look upon it as that. When I said at lunch time I enjoy, not necessarily company, but good company, a question arose as to what did I think of you and your secretary. And my reaction then, as I recall it, was that these people are friends. Not that these are people that I have to please, that I'm working for. These are people that I owe a debt to, such as parents, these people are friends. As much as some of the people with whom I work and whom I call friends.

Dr. But still you were in this web of the experiment and of therapy!

Mr. K. Well, this is a web of my own choosing.

Dr. So is the web of your marriage.

Mr. K. Yes, and I think that was chosen from lack of experience and lack of knowledge of people and a fear of being alone. I think if you listened to my record of last week—

Dr. I don't think I've seen it transcribed as yet.

Mr. K. Well, in one part there I said the reason I think I married my wife is that she was the first girl that had returned my love and affection, and that I married because I was afraid of losing her forever. If I had gone to Seattle to work as I did immediately when I got out of college—this was when we were married—that if I had left at the time without marrying her, I would not have seen her again in all probability, and she would have been lost to me. I think this is the deep seated reason that I got married at that time. I was afraid of losing the first person who had loved me.

Dr. You didn't want to be alone in that web anymore?

Mr. K. It might have been that. I think I wanted to be free of that web. You see, when I say this other personality of mine is watching what I'm doing and counting to see how far I can lose my temper, this is the one that's up in the corner of the web, watching the people fight down below—that's exactly the same feeling. Actually, it's a part of me up at the top, watching myself fighting with myself down below.

Dr. Then you possibly could be the spider in the dream?

Mr. K. We're not spiders. We're on a spider web, but we're people.

Dr. Yes, but you also describe the spider as a praying mantis.

Mr. K. No. I was searching for a descriptive term there.

Dr. How do you feel now? How are your eyes?

Mr. K. I can see things fairly well. I still notice effects of the drug in my eyes. I'm feeling warm. Some of the tension has returned to my arms and legs—not so much. I'm still breathing rather fast.

Dr. I think we'll modify your reaction with phenobarbital. Rest your eyes.

Although the emphasis has been placed upon ego reinforcement it must be understood that the reinforcement is not only due to the processing by the unconscious of the material supplied by the drug, but also to the activity of the therapist. In order to emphasize the therapeutic relationships, excerpts from the same eight-hour experiment of *K* during which therapy was not emphasized or stressed show that ego depression can occur with the same dosage of 50 micrograms.

Subject—*Mr. K.*

August 1, 1952

LSD-25 (orally)

50 micrograms

Mr. K. . . . I find that I am annoyed by something. In putting these cards back into the box, the secretary has stacked them so that the cards with a similar cut on the corner have been stacked together, possibly two or three hundred at a time. And in going through these questions, I find I'm really annoyed by the fact that I've perhaps answered fifty or

one hundred in a row 'true' and the next fifty or one hundred 'false'. So I'm now taking the cards out and shuffling them up a bit to make them more interesting.

Dr. Why were you annoyed?

Mr. K. I don't know why I was annoyed, but I know that I was annoyed enough that I wanted to change them by shuffling them thoroughly. This brings something to mind. You mentioned that I might resent authority. Of course, it might be that I resent making the same choice that somebody else has previously determined for a different reason.

Dr. Did that feeling of annoyance come on you suddenly?

Mr. K. No. When you say suddenly that involves a time sense. It seemed to be building up, but the building up time was short. It didn't appear to be suddenly. A little while ago I was sitting in the chair and feeling a great lassitude. The thought came to me, to describe how it felt to move. If I relax and do nothing but only think about the feeling, I become very conscious of myself. I feel very reluctant to do anything. Reluctant to move! To think or to read a magazine! I just want to sit. I just want to sit and be conscious of myself. I think that if I had a more comfortable chair, I would sink about four feet deep in it. Yet, at the same time that I had the feeling that I've just expressed, a small part of my mind is quite restless, and I want to read a magazine, if I could summon the physical energy to do it. I'm wondering how fast time is passing for me. Every time I take a distinct memory of everything I'm doing and saying and the peculiar instant that I'm living in, is very, I don't know. It gets rather vague. My time sense diminishes. If I try to figure something ten minutes back, I'll probably be only about sixty per cent accurate. If I try to figure something five minutes back I'm only ninety per cent accurate. If I try to think of something within one second of this particular moment, as I did a second ago, of what I will do in the next second, it is very real to me, and I think in the appropriate time sense. Whereas, if I get anything fairly away from that particular instant, I get very confused. If I had to do something five minutes from now, I don't know if I would do it. I'm feeling very depressed. I have a much lower hope for the world than generally.

Dr. Yet, you smiled when you said, "I'm feeling very depressed." You smiled broadly.

Mr. K. I think that at the same time I'm talking I have an inferiority feeling. What does one man's opinion matter! I feel that I'll never make any earth-shaking changes or have any pronounced effect on anything. You know how it is. You sometimes make a profound statement, yet smiling, just trying to sort of disarm your statement. I don't know why people do it, but they do. You try to emphasize a point and at the same time to de-emphasize it.

Dr. How about your attitude toward authority now?

Mr. K. At the moment and at the moment only, I think I'm very submissive to authority. I might mentally fight it, but I think I would go along with anything that anybody told me to do.

One week later when Mr. K took 50 micrograms of LSD-25 orally, he fell asleep in the office waiting room and had a dream in which he was underground in a tunnel. A study of the verbatim recording of this dream which follows, illustrates how the process of ego reinforcement has continued in Mr. K, how he rapidly developed insight into his adaptive mechanisms, and how well he was able to understand what was happening to him under the influence of LSD-25.

Subject—Mr. K.

August 8, 1952

LSD-25 (orally)

50 micrograms

Mr. K. I had a dream. The duration of the dream was only about two or three minutes.

IT SEEMED I WAS UNDERGROUND IN A TUNNEL PERHAPS SIX FEET HIGH AND SIX FEET IN DIAMETER. IT WAS DARK, DIRTY, AND POORLY LIGHTED. THERE WERE SOME PEOPLE WITH ME. THEY WERE MEN AND WE WERE BREAKING OUR WAY OUT OF PRISON. WE HAD JUST DRILLED A HOLE IN THE ROCK FACE OF THE TUNNEL WITH A HAMMER AND CHISEL, AND WE HAD JUST LOADED BLASTING CAPS INTO IT. WE HAD JUST FINISHED LOADING THE BLASTING CAPS INTO IT WHEN I HEARD YOU SAY, "ALL RIGHT, YOU CAN GO BACK TO YOUR ROOM NOW."

That was the full extent of the dream. It was very short. I was standing in the background. I wasn't actively participating in the dream. In other words, I hadn't been swinging the hammer. I just seemed to be present. It didn't affect me whether we broke out of prison or not. I wasn't looking for the guards and I wasn't looking to break out. I was a spectator. I wonder if I am always a spectator. I wonder if I'm a spectator in watching other people do things, unless I am working alone. That is the only time I can work, when I am alone. I'm free associating now.

Dr. That's all right. Go right ahead.

Mr. K. It suddenly occurred to me that whenever there are discussions or talks about technical matters, I prefer to listen to the opinions of others. I form my own opinions but I don't voice them. I merely form them for my own reference. And I think that it is only under direct stimulus that—under a direct question, or if somebody definitely asks my opinion on something, well, I give it and then I have no trouble giving it. If I have an opinion I express it all right. I get the point over, but I won't do it unless I'm invited to. It's very—not rarely, but I think it's unnatural with me to, in a formal group, to spontaneously do anything by myself. Although I can remember instances where I have. I went to a town meeting, about a variance in a building code which I had been asked for by a neighbor, and I remember spontaneously getting up and saying what I thought about the matter in front of people. There were perhaps forty or fifty people who were total strangers to me. I knew nobody there. In general, in everyday life, in business, I will not venture

an opinion unless it's asked for. But this is only in formal meetings, particularly where there are superiors present. With my—I find it easier—I find it easy to associate with those men who are working under my supervision. I feel able to get close to them. In other words, if there were a superior and inferior and if I am the superior, I feel it very easy to break down any barriers and become very free and easy to associate with the men who are working for me. I suppose in this way I am always identifying myself with the underdog. I wonder if I am the one who feels discriminated against, but only in this battle between myself and my superiors, or myself and authority. And yet, if there is always a battle between myself and authority, it is never a battle between people working under my supervision. I'm never conscious of taking or being on the side of authority. I would invite some questions or leading remarks at this time, doctor.

Dr. You don't need any. Go right ahead.

Mr. K. I always have the feeling that I am working for the men under me; that I am working for the men that I'm responsible for. In other words, I have a foreman and anywhere between three and fifteen men working for me on technical work. I feel that I am a buffer for them, between them, and I'll use the word again—authority—that I should protect them from changes in managerial opinions from week to week; that I should protect them from the worries of budgets; from the worries of whether we're going to run out of work or have too much work coming up. I feel I should always work for them and see that they have the proper instructions and tools and the proper safety equipment. I feel that the work should be outlined for them and yet, on the other hand, I know that this is seldom done for me, and I realize that it doesn't have to be done for me. I mean, I say it now, as a logical opinion; logic tells me that this work doesn't have to be done for me, but somehow I resent it. I have the feeling that I do more for other people than they do for me. I think of people under me more carefully than the demands that are put upon me. I feel I am more reasonable than the people above me.

Dr. In the dream, you were underground?

Mr. K. Oh, yes, in this dream. I started quite a long discussion there about why I was a spectator.

Dr. You were brought up that children are to be seen and not heard.

Mr. K. Yes, but this was a long time and a lot of experiences ago. I can't see that—I'm not too sure that that would be exerting a strong influence on me. But something definitely is exerting a strong influence.

Dr. That you should be seen and not heard. Remember that was what your boss told you, that you were not heard enough? So that means that you are adapting yourself to life by being seen and not heard.

Mr. K. I've been trying to adapt myself the other way this last week and really have been giving him the treatment. Now why did I put it that way—giving him the treatment? I've been asking him every ques-

tion that came into my mind and he seemed to like it. He liked it very much. He told me to take the day off. I have the feeling, though, that I can't stand it too long, because I know that I'm not sincere about it.

Dr. Well, don't you think that there is an intermediate position?

Mr. K. Yes, there must be, but at the moment I don't. I feel antagonistic toward him, very antagonistic. The thing to do is not to find the intermediate position, but to want to find it. I would feel satisfied if I did find it. I just can't believe, somehow, that being a brain picker as he is, and a seeker after other people's ideas, that that is the right way to do things. The right way to get a job done, that is, until I have been convinced, or have seen, or something has been proven to me that there is a better way of doing it—I find it very hard to camouflage my feelings, to operate this way.

Dr. It may not be the best way, but under the circumstances it may be one of the effective ways. It all depends on what you mean by the word 'best'.

Mr. K. Momentarily, I mean for a matter of days, when he explains to me that the spreading out of information was what the laboratory wanted, I was convinced. I wanted to be convinced and I believed it and I felt it for a short time, but the old habit of working seemed to come back again and I can't get on to this new area of spreading ideas far and thin. I guess I tend to work by taking all the ideas and concentrating them into a single line or a small area so that they will be effective. I get them down on paper to get them made into metal—to get them produced. I don't think I'm essentially a talker or perhaps I am a talker and never developed the habit.

Dr. You're doing very well now.

Mr. K. I think a lot to myself.

Dr. In discussing your dream you said that your friends were blowing up things under you.

Mr. K. Yes. We were digging this tunnel, trying to blow our way out of this prison we were in. We were trying to escape; trying to get away from it. No! They were trying to get away from it. I was there. I seemed to be taking no active part in it. The duration of the dream was only three or four minutes.

Dr. You don't think that you could have been the other people in the dream?

Mr. K. No.

Dr. Who were they?

Mr. K. I don't know. Total strangers. I didn't know them. I hadn't known them in the dream. Yes, I did too! In the dream I had known them. I had been associated with them. I don't know how much. I knew nothing about them. I didn't know their names. I didn't know what they looked like nor how they felt.

The integrative processes which occurred with Mr. K slowly developed to the point where he was able to utilize in his daily life the insight which

he developed through these and other psychotherapeutic interviews. At present he has an executive position in which he is both seen and heard, his salary has increased more than 50 per cent, and the adaptive processes are better than they have ever been.

D. DISCUSSION

1. *Transference*

In the foregoing verbatim recordings there were no subjects or patients undergoing prolonged psychoanalysis. Psychoanalytically oriented brief therapy was the technique employed by the writer. The results of the use of LSD-25 in investigations with more formal psychoanalytic procedures has, in three cases studied, provided a technique of facilitating the elimination of psychic blocks related to the transference. Both unconscious and preconscious negative and/or sexual guilt provoking attitudes toward the therapist have emerged so that more rapid progress in reconstructive therapy was feasible. For example, during the two hundred and fiftieth interview of Mr. L under 40 micrograms of LSD-25 taken orally, the stationary state of the analysis was favorably modified by the patient's understanding his deep distrust of the therapist. This distrust had prevented the patient from disclosing his feelings toward the therapist, much less understanding them.

Dr. Now, to go back to your dream. Do you feel the fact that you've converted me into an oriental tyrant. . . .

Mr. L. (Interrupts) With a smiling face.

Dr. . . . means that you feel that my attitudes are unjustified?

Mr. L. Well, I know what it really means now; *I didn't until this very second, I think*. That is, just now, as you were speaking, it may have been wrong of me, but I was listening to you with half of my mind, and the other half of my mind was saying, "Now, just what the hell did I mean by the 'oriental'—what is meant by the 'oriental'," and then suddenly the word '*distrust*' came out! And there I am doing the very thing I condemn and hate so much in other people. I can realize now why the feelings are so strong, because they some way reflect or blanket my own feelings to a degree. Ah, now, not trusting you is also another way of not trusting what I, or we, may find in what we uncover as we go along in our work, in what I say, in what you say, and so on. In other words, I don't fear you! *I fear what you may uncover*. Or, *I fear what we may uncover in myself*. It's actually saying the same words again and again. So the classroom scene is so clearcut. I wonder if that's why I asked if there's such a thing as being too clearcut. You know, it was so obvious it meant something else, but I doubt it. It doesn't, in light of what we have been discussing, the plan, using LSD and so on. And yet, I'd like to know why I've been so anxious to use LSD.

You know, if you can remember, I sort of have been hinting at using this for, well, for quite some time. I've been very anxious to use it.

The foregoing excerpt is taken from 33 pages of a single spaced type-written recording of a four hour interview under the influence of LSD-25. Although the patient had previously hinted that he distrusted me, the brilliant manifestation of his underlying fear of me was utilized effectively to develop insight into his homosexual fears of his father. The details of this interview will be reported elsewhere.

2. *Group Therapy Possibilities in Non-Psychotic Patients*

Many of our experiments were performed in groups of two to five subjects who were not patients but were volunteers paid \$20 per day. When the same subjects (non-psychotic) met repeatedly, irrespective of the dose up to 100 micrograms, a relationship developed amongst them which has apparently *always* led to some insight and better adaptive techniques both at work and in community activities on the part of the subjects. It is realized that a basic need for psychotherapy may have led these subjects to repeatedly volunteer. But the problem here is not the motivation for taking LSD-25 but rather the results of taking it. The production of euphoria in doses between 25 and 50 micrograms gives this dose range (preferably 25 micrograms, orally) unique properties (4). The therapist or group leader often need only catalyze the activity of the group. He may even withdraw from the room and permit the group to function under his control, but without his direct leadership. Surrogate leadership rapidly develops. It is of interest that acting out of hostilities by subjects within the group does not lead to the loss of their social status within the group. I believe that this is different from group responses if alcohol were the euphoriant. LSD-25 then, assumes a directing force, which through modification of the ego as discussed in the foregoing, results in a therapeutic process. The disadvantage in the use of LSD-25 as an adjuvant is the length of the session which may require three to four hours. However, with doses of 25 micrograms, no important changes occur for about 30 to 60 minutes. The group leader may, therefore, safely begin 40 to 60 minutes after the administration of the drug and discontinue at about three-and-a-half hours after administration of the drug. Precautions have been previously described (1).

3. *Effect on Reassociation of Dreams* (2)

I have previously reported that earlier concepts and techniques of Freud may be amplified by the use of verbatim recordings of dreams.

The repetitive dream and its significance have been of importance in theory and in therapy. The command dream following hypnosis has also been the subject of considerable investigation. By commanding the patient under light hypnosis to reassociate a dream already recorded verbatim, ego development occurring during psychoanalysis may be utilized at later stages of therapy. The verbatim transcripts of the first dream and three subsequent reassociations of the same dream were the basis of the previous study reported. The patient is commanded to redream the dream which is repeated verbatim to the patient who then relates the dream as if he had dreamed it, while awakening from light hypnosis under which the dream was "redreamed."

Dream reassociation offers several advantages: (a) The patient may be regressed by the therapist at will to an earlier stage of the analysis. (b) The negative transference at times may be readily investigated. (c) As ego development proceeds, unconscious material previously threatening to the patient may be more freely discussed and evaluated.

It must be emphasized that the therapist's notes cannot take the place of the verbatim technique of dream reassociation. Each word expressed by the patient represents almost an infinite number of symbolic processes. The introduction of a word not used by the patient will alter the process of dream reassociation and contaminate the procedure by the therapist's screened summary of the dream. Verbatim records are of unique importance in the use of LSD-25 which I have employed as an adjuvant to facilitate dream reassociation. The same dream is reassociated without and then with LSD-25. Verbatim data of this type have been obtained and will be the subject of future communications. It is my present belief that LSD-25 offers a safe and effective way of simultaneously producing sufficient ego depression to enable new unconscious material to enter consciousness, and sufficient ego enhancement to enable the patient to make rapid and effective use of the insight developed so that reconstructive processes may be expedited.

4. Length of Time Required for Therapeutic Interviews

The "50-minute hour" is not adequate if the therapist desires to use LSD-25. With therapeutic doses of LSD-25 between 25 and 50 micrograms it is desirable to keep the patient in the office at least five hours after the ingestion of the drug. The patient, knowing that there is almost limitless time compared with the ordinary 50-minute hour is able to face the consequences of dealing with preconscious and unconscious material that hitherto could not be readily tolerated. Several patients after four-hour LSD-25 interviews, desired planned sessions lasting two hours without LSD-25. It is

the purpose of psychotherapy not to provoke anxiety, but rather to utilize the anxiety induced by psychotherapy to aid the ego to adapt to the emergence of preconscious and unconscious cathexes and hypercathexes. It is my belief that LSD-25 in the hands of the therapist who understands the nature of the psychopharmacologic action of the drug can organize the therapeutic interview to achieve this desirable goal of ego reconstruction.

E. SUMMARY

1. Previous data on the nature of the ego enhancement occurring during the LSD-25 reaction is amplified by making a direct comparison with the ego depression simultaneously occurring. During the LSD-25 reaction advantage may be taken of the integrative function of the ego if the therapist recognizes the presence of the processes of ego reinforcement.
2. Although ego depression may occur during the LSD-25 reaction to produce a psychotic state, this state is associated primarily with doses greater than 50 micrograms by mouth. The nature of the ego depression is shown by a verbatim recording. The subject was essentially incapable of communicating effectively with the interviewer during the height of the LSD-25 response.
3. The nature of the integrative processes during ego enhancement with small doses of LSD-25 (20 to 50 micrograms) is illustrated by verbatim recordings. It is emphasized that during the LSD-25 reaction, the integrative functions of the ego often operate effectively to utilize preconscious and unconscious material during therapeutic interviews lasting as long as four hours and covering periods of observation lasting six hours. Whether or not the LSD-25 reaction develops into one where ego reinforcement or one where ego depression is emphasized depends to a great extent upon the relationship of the therapist to the patient. During the same therapeutic interview the therapist can manipulate the nature of the responses.
4. It is believed that LSD-25 may be utilized to more effectively mobilize psychodynamic vectors during the analysis of the transference.
5. In non-psychotic groups studied, important relationships developed amongst groups consisting of two to five members. These always led to some insight and better adaptive techniques both at work and in community activities on the part of the subjects who met repeatedly. This suggests the possible use of LSD-25 as an adjuvant to group therapy.
6. Preliminary data are reported on the effect of LSD-25 on the re-association of dreams. It is stressed that verbatim recordings are a necessary condition for the successful utilization of dream reassociation.

REFERENCES

1. ABRAMSON, H. A. Lysergic acid diethylamide (LSD-25): III. As an adjunct to psychotherapy with elimination of fear of homosexuality. *J. of Psych.*, 1955, **39**, 127-155.
2. ———. Reassociation of dreams. *Am. J. Psychiat.*, 1955, **112**, p. 300.
3. ABRAMSON, H. A., JARVIK, M. E., GORIN, M. H., & HIRSCH, M. W. Lysergic acid diethylamide (LSD-25): Tolerance development and its relationship to a theory of psychosis. *J. of Psych.* (in press).
4. ABRAMSON, H. A., KORNETSKY, C., JARVIK, M. E., KAUFMAN, M. R., & FERGUSON, M. W. Lysergic acid diethylamide (LSD-25): XI. Content analysis of clinical reactions. *J. of Psych.*, 1955, **40**, 53-60.
5. BUSCH, A. K., & JOHNSON, W. C. LSD-25 as an aid in psychotherapy. *Dis. Nerv. System*, 1950, **11**, 241-243.
6. FREDERKING, W. Ueber die Verwendung von Rauschdrogen (Meskalin und Lysergsaurediethylamid) in der Psychotherapie. *Psyche*, 1953-54, **6**, 342-364.
7. SANDISON, R. A. Psychological aspects of the LSD treatment of the neuroses. *J. Ment. Sci.*, 1954, **100**, 508-515.
8. SAVAGE, C. Lysergic acid diethylamide (LSD-25). A clinical-psychological study. *Amer. J. Psychiat.*, 1952, **108**, 896-900.
9. STOLL, W. A. D-lysergic acid diethylamide (LSD-25), a phantasticum of the ergot group. *Schweiz. Arch. f. Neurol. u. Psychiat.*, 1947, **60**, 1-45.
10. ———. Ein neues, in sehr kleinen Mengen wirksames Phantastikum. *Schweiz. Arch. Neurol.*, 1949, **64**, 483-484.
11. WEYL, B. Versuch einer psychopathologischen Analyse der LSD-Wirkung. *Diss. Freiburg i. Br.*, 1951.

133 East 58 Street
New York City 22

A NOTE ON THE NORMAL PROBABILITY FUNCTION IN COLOR VISION*

Department of Psychology, University of Stockholm

GÖSTA EKMAN¹

A. INTRODUCTION AND PROBLEM

In a previous paper (1) it was shown that five color factors were necessary to account for the matrix of subjective similarities between spectral colors. When the factor loadings were plotted against wave length, they closely resembled the curves of relative excitability of nervous elements in the retina, which have been obtained by Granit (2, 3).

After the previous analysis was published, a small adjustment of two of the axes (R and Y) was performed, in order to eliminate a slight skewness of the Y curve. This transformation is shown in Table 1. The orthogonal rotated factor matrix after this adjustment is given in Table 2. The entries of this table are reproduced by circles in Figure 1.

It is the purpose of this paper to show that our five color factors may be adequately represented by normal curves.

B. FITTING NORMAL CURVES

Since the entries of Table 2 are factor loadings which are not equally spaced along the base line, the routine procedures of fitting a normal curve to a frequency table are not applicable. We had to resort to a less formal procedure. Free-hand curves were drawn to connect the points of any factor on the graph, and ordinates were interpolated with intervals of five millimicrons. From these ordinates, and the interpolated location of the maximum, constants of the normal probability function were computed. They are given in Table 3.

The normal curves are drawn in Figure 1. By inspection, the correspondence between theoretical curves and factor loadings may be judged as rather good. It should be noted that the method employed in fitting the curves does not provide the best possible fit. In the least squares sense, we

*Received in the Editorial Office on October 15, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹This investigation was made possible by a grant from the Magnus Bergvall Foundation. The technical work was carried out by Mr. B. Dahlbäck.

have at hand only a conservative estimate of the fit. It may also be pointed out that further slight adjustments of the factor matrix would probably increase the agreement between theoretical and analytical data.

TABLE 1
ADDITIONAL TRANSFORMATION MATRIX

	Y_{12}	R_{12}
Y_1	.985	.174
R_{12}	-.174	.985

TABLE 2
ADJUSTED ROTATED ORTHOGONAL FACTOR MATRIX

Wave length	V	B	G	Y	R
434	.913	.128	-.030	.031	.091
445	.904	.188	-.009	.006	.082
465	.385	.746	-.030	.023	.010
472	.341	.804	.004	-.001	.006
490	.109	.678	.364	-.036	.003
504	.012	.381	.666	-.046	.011
537	.095	.074	.836	.011	-.008
515	.086	-.006	.810	.088	-.003
584	.003	-.020	.320	.585	.209
600	.013	-.004	.146	.792	.379
610	.004	.028	-.002	.525	.687
628	.038	.000	.069	.210	.883
651	.054	.008	.014	.021	.927
674	.093	-.021	.028	-.007	.830

TABLE 3
CONSTANTS OF THE NORMAL PROBABILITY FUNCTION

	V	B	G	Y	R
Area	(10.18)	8.17	16.36	7.21	(19.22)
Mean	438.00	475.12	536.55	596.64	650.00
SD	23.02	19.84	35.63	19.45	39.00

The goodness of fit has been tested in another way as well. A similarity matrix has been computed from the factor matrix of Table 2. It has been compared with the empirical similarity matrix, and the correlation between the 91 pairs of values is shown in the left hand graph of Figure 2. The right hand graph shows the correspondence between the same empirical similarity matrix and the one which is inferred from the normal curves of Figure 1. Again the agreement is good, the correlation coefficient being $+.98$, as compared with $+.99$ in the former case.

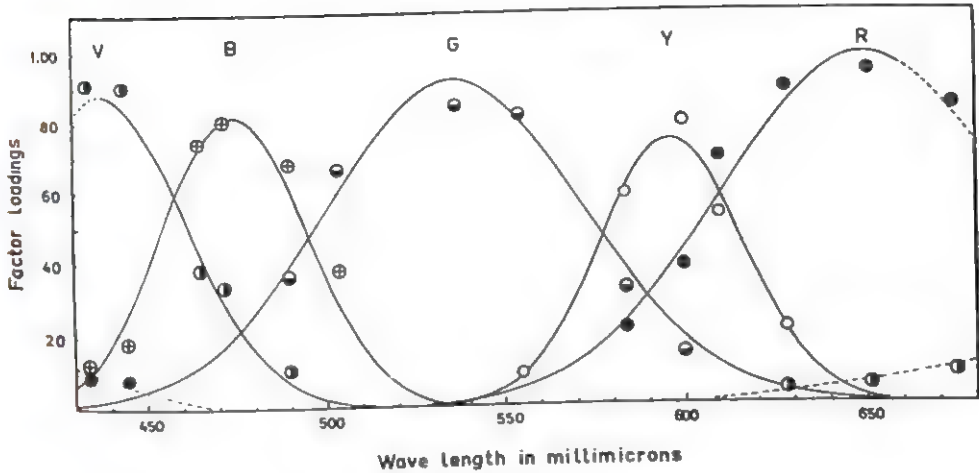


FIGURE 1

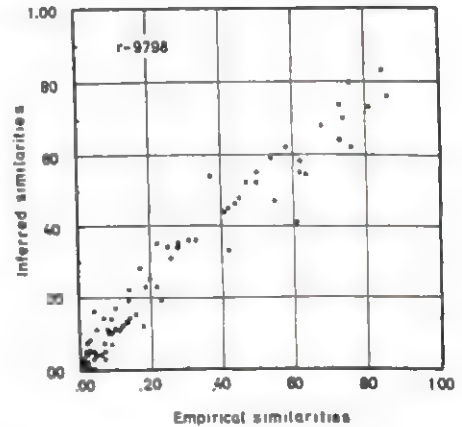
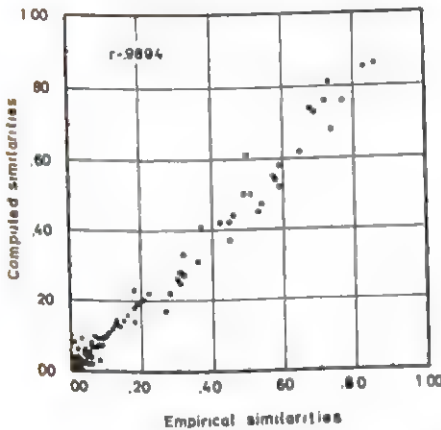


FIGURE 2

C. CONCLUSIONS

In view of the above considerations it may be stated that our five color factors are fairly well represented by normal curves.

This statement has some interesting implications. Any one of the five color factors is a function of wave length,

$$y_1 = f_1(\lambda),$$

and the color (hue) is a function of the ordinates at the spectral point, representing an experience vector in the five-dimensional subjective space,

$$C = \phi(y_1, y_2, \dots, y_5).$$

Hence, in principle, the color is a function of wave length,

$$C = \psi(\lambda).$$

The constants of this function may be considered attributes of the sensory and nervous system which transforms the unidimensional wave length of visible light into a multidimensional variation which is experienced as qualitative. These constants are to be determined with much more accuracy than our present estimates which are, so far, based only on preliminary experiments.

REFERENCES

1. EKMAN, G. Dimensions of color vision. *J. of Psychol.*, 1954, **38**, 467-474.
2. GRANIT, R. *Sensory Mechanisms of the Retina*. Oxford: Univ. Press, 1947.
3. ———. *Receptors and Sensory Perception*. New Haven: Yale Univ. Press, 1955.

*Department of Psychology
University of Stockholm
Teknologgatan 8
Stockholm, Sweden*

THE INFLUENCE OF AUDITORY STIMULATION ON ANISEIKONIC PERCEPTION: I. A PRELIMINARY STUDY*¹

The Menninger Foundation

JAMES E. SIMPSON²

A. RATIONALE

The rationale of this preliminary investigation of the influences of pleasant and unpleasant sounds on aniseikonic perception stems from several areas of research. Murphy and others have been long concerned with autistic effects in perception and, more specifically, with the rôle of positive and negative affect in perceptual learning (3, 4, 5). For example, in figure-ground studies with adult Ss the perceptual dominance of the stimuli associated with pleasant affect and the subordination of those associated with unpleasant affect have been variously demonstrated in several laboratories (8, 10, 11). The possibility that autistic effects could be experimentally controlled in natural and more complex visual fields—a global area like a room rather than a fractional area usually employed in figure-ground experiments—was viewed as a question of the influences of affect on perception in everyday life situations. Observations of ourselves and others commonly indicate the everyday occurrence of the influence of affect on perception of complex visual stimuli; indeed, a disputed touchdown, an earthy Broadway play, or a political leader are rarely perceived without affective influences. These kinds of stimulus fields are perhaps too complex for laboratory control, but some of the positive and negative components of affect are identifiable and the autistic perceptions can be predicted and verified (6). Similarly, a simple autism theory model would predict the perceptual motions in the directions of the positive and negative affective values of experimental stimulus fields. For example, if two identical rooms are seen by Ss under identical conditions except that Room *P* is experienced with pleasant affect and Room *UP* with unpleasant

*Recommended for publication by Gardner Murphy, received in the Editorial Office on October 24, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹In addition to the support of The Menninger Foundation this investigation was supported by a research grant (M-715C) from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service.

²The author is grateful to Dr. Gardner Murphy for his continued encouragement and guidance and to Drs. Julian E. Hochberg and Wayne H. Holtzman for their positive aid with some negative statistics.

affect, the autism theory model would predict subsequent positive approach or acceptance behavior to Room *P* and negative resistance or rejection behavior to Room *UP*. In the specific hypotheses detailed in Section C below, the positive accepting perceptual responses to Room *P* and the negative resisting perceptual responses to Room *UP* are primarily described in terms of the *decrease* in the time barrier (reaction time) for the perception of the pleasant-associated visual field compared to the time barrier for the perception of the unpleasant-associated visual field.

B. EXPERIMENTAL VEHICLE

To test this mutability of human perception to such simple affective associations, aniseikonic lenses were selected as the experimental vehicle because this device provides a visual experience that is relatively unrelated to previous experience and is unique for each *S*—as Ames said, "It must be emphasized here that the nature of the phenomena [aniseikonic distortion] . . . i.e., the way things look, can be thoroughly appreciated only if the individual actually experiences them (1, p. 338)." The research of Ames, Cantril, and associates has shown the wide variations of aniseikonic perception in individuals and, more strikingly, in different visual fields (1, 2, 9, 12). It seems that the mutability of the aniseikonic effects is dependent on subtle psychological variables as well as physical attributes of the perception matrix. For examples, a normal room will not "distort" as much as the unstructured "Leaf Room" (all surfaces leaf-covered) (7, p. 54, Figure 14.3); still water will not tilt but the adjacent flat lawn will. Of the latter example, Ames states: "An explanation of this phenomenon must be based on the fact that *O* knows from past experience that grass surfaces can be inclined but that the surfaces of bodies of water are always horizontal (1, p. 349)." In the former example of normal room vs. leaf room the aniseikonic effects seem to be mainly determined by the physical structure of the visual field. The normal room provides many monocular cues which tend to dominate the perception and resist aniseikonic distortion, whereas the binocular disparity built into the aniseikonic spectacles comes into dominance in the leaf room where monocular cues are minimal. So if the optical structure of the aniseikonic lenses and the binocular-monocular cues in the visual field were to be experimentally controlled, this mutability could be channeled to test the psychological, rather than the physical, factors in aniseikonic perception.

C. HYPOTHESES

The psychological factors to be tested were the rôles of affect in changing aniseikonic perception, or, in less general terms, the influences of pleas-

aniseikonic effects are: (a) the backward slant and increase in size of one side of the backwall of the room; e.g., with the "Right Lens" (right lens magnifies and left neutral) the backwall appears trapezoidal, largest at the right corner, and slanted backward left to right about 20 degrees so that the right side paradoxically appears both larger and farther away; and (b) the induced distortion is *not* seen immediately; i.e., the room appears normal for several seconds or minutes, and then the distortion occurs suddenly in full magnitude. In this straw room considerable variation occurs between Ss for these two main effects as well as for the other aniseikonic effects of markedly tilting floor and slanting sidewalls and ceiling.

3. *Experimental Design: Independent Variables in Three Pairs*

A factorial design was employed for the repeated measures of 28 secretarial students divided into four groups of seven girls; i.e., each of the seven Ss in any one group was trained and tested with the same combination of the six independent paired variables. These independent variables are: (a) Affect: pleasant music vs. unpleasant noise: *M* vs. *N*. (b) Lenses: right spectacles (backwall slants back on right side) vs. left spectacles (backwall slants back on left side): *R* vs. *L*. (c) Stimulus-Order: also only two possibilities; either *S* is presented with music first and noise second or vice versa: *1st* vs. *2nd*. The four possible combinations of these independent paired variables are the four order sequences (using the above symbols): Order I, RMLN; Order II, LNRM; Order III, LMRN; Order IV, RNLN.

4. *Measurement of the Dependent Variables: Reaction Time and Distortion*

The Ss were not aware of the purpose of the experiment and believed it a proficiency test. Each *S* was told: "This is a test to measure your accuracy under different conditions." These conditions were five training trials with constant music and five training trials with constant noise followed by the same number of testing trials in silence with *S* wearing the spectacles in the same order as in the training trials; i.e., the experimental routine was an "a:b:a:b" sequence. The separate scorings of the seconds of reaction time and of the degrees of distortion were made by the subtractive method to indicate the percentage change between training and testing trials with the lenses associated with music as against the change with the opposite lenses associated with noise. The differential change in reaction time associated with music vs. that with noise is designated as "reaction time shift" and similarly for the change in degrees of distortion, "distortion shift." In short, a reaction time shift that was *positive* (i.e., supports the hypothesis) occurred

when *S*'s reaction time for music-associated distortion *decreased* relative to the reaction time for noise-associated distortion; similarly, a *positive* distortion shift occurred when the degrees of distortion perceived in association with music *increased* relative to the degrees of distortion associated with noise.

5. *Assessment of the Intervening Variable: Affective Potency of the Auditory Stimuli*

The affective responses to the music and to the noise were assessed for each *S* in a post-test interview. The intensity and duration of the pleasantness of the music and of the unpleasantness of the noise were rated separately by *E* for each *S* and then ranked for magnitude of induced affect. The mean rank of the noise and music rankings gives an index of the affective potency of the auditory stimuli for each *S*. The ratings made by *E* were based mainly on the interview report of each *S*, but *S*'s spontaneous evaluations verbalized during the training trials were also considered.

E. RESULTS AND CONCLUSIONS

The over-all results show that 25 of the 28 *Ss* gave a positive reaction time shift and/or a positive distortion shift; i.e., they gave support to the primary hypothesis of reaction time shift and/or the secondary hypothesis of distortion shift. Further, 9 *Ss* gave positive results for both reaction time and distortion shifts, and no *Ss* gave negative results for both shifts. Although this combination of the reaction time results and the distortion results show the general trend to be positive, these two separate measures of the aniseikonic effects were analyzed separately as appropriate to primary and secondary hypotheses.

The primary hypothesis of reaction time shift—that distortion experienced with pleasant affect will be accepted more quickly than distortion experienced with unpleasant affect—was supported by the reaction time data. There were 17 *Ss* who gave positive reaction time shifts, 7 *Ss* gave negative shifts, and 4 *Ss* gave neutral shifts. These results are not without statistical significance (chi-square: 4.17; $P < .05$). The secondary hypothesis of distortion shift was not supported by the results (17 *Ss* positive; 11 *Ss* negative). The complex data of this experiment cannot be clearly analyzed because the confounded factorial design does not allow for separation of the counterbalanced paired variables; e.g., *F* tests of the distortion data for Music \times Noise and First \times Second stimulus-order are equally significant

(means of 19 & 5; $P: < .03$) but for perfectly confounded, unseparated effects. Also, the reaction time data per se present difficulties: (a) the extreme individual differences (e.g., means of *Ss* range from 0 seconds to 282 seconds; training to test changes range from 0 per cent to 7,068 per cent) and (b) three of the four *Ss* who gave neutral reaction time shifts showed zero seconds reaction time throughout the training and test trials so probably should not have been included in the sample. Further complexity is found in the interrelations with the intervening variable, the affective potency of the auditory stimulation. This measure of the degree of affective response to the pleasant music and the unpleasant noise (as ranked by *E* according to post-test interviewing of each *S*) is apparently unrelated to order of presentation of the music and noise (CR 1.01; $P: .31$, two-tailed) but appears to be differentially related to reaction time and distortion shifts. For example, the 9 *Ss* of lowest rank for degree of affective response were among the highest for reaction time shift (8 positive; none negative; one neutral) and among the lowest for distortion shift (3 positive; 6 negative). Further, the 9 *Ss* of highest rank for degree of affective response showed opposite shifts: lowest for reaction time shift (3 positive; 5 negative; one neutral) and highest for distortion shift (7 positive; 2 negative). The data of this preliminary experiment indicate differential mutability of aniseikonic perception to pleasant and unpleasant affective influences. However, because the confounded factorial design employed does not allow for separation of the interesting interrelations of the delicate aniseikonic effects of reaction time and distortion and the degree of affective responses, further experimentation of less complex design is indicated and so planned.

The planned experiment will exclude the confounding of variables by using only one stimulus with only one kind of aniseikonic spectacles for each of three groups of *Ss* in three series of trials. For the two experimental groups the first series will be silent pre-training trials for base-line data; the second series will be stimulus-training trials (with music for one group and noise for the other) for stimulus-response data; and the third series will be silent testing trials for perceptual learning data. The third group will have three series of silent trials—without the experimental variables of pleasant or unpleasant sounds—for control data. This simple design points more directly toward the influence of induced affect on perception but without the contrasting of the pleasant and unpleasant for each *S* as in the preliminary experiment; so the differential responses will not be measured in the proposed experiment and must await another experiment.

F. SUMMARY

By testing 28 young women who were given a pseudo-test of proficiency in adjusting a selsyn-controlled measuring-arm while wearing aniseikonic spectacles, the training vs. test changes in reaction time and degrees of distortion that occurred in association with pleasant music and unpleasant noise were obtained. The reaction time data show that generally the aniseikonic distortion previously experienced with pleasant music was more quickly perceived than the distortion previously experienced with unpleasant noise. The secondary measures of the amount of distortion are not statistically significant in showing differential aniseikonic responses to the pleasant and unpleasant sounds. Because the confounded factorial design employed in this preliminary experiment does not allow separation of the interrelations of the delicate aniseikonic effects and the degree of affective influences, further experimentation of less complex design is indicated and so planned.

REFERENCES

1. AMES, A., JR. Binocular vision as affected by relations between uniocular stimulus-patterns in commonplace environments. *Amer. J. Psychol.*, 1946, **59**, 333-357.
2. ———. An interpretative manual for the demonstrations in the Psychological Research Center, Princeton University: the nature of our perceptions, prehensions, and behavior. Princeton: Princeton Univ. Press, 1955.
3. MURPHY, G. Personality: A Biosocial Approach to Origins and Structure. New York: Harper, 1947.
4. ———. Affect and perceptual learning. *Psychol. Rev.*, 1956, **63**, in press.
5. ———, & HOCHBERG, J. E. Perceptual development: some tentative hypotheses. *Psychol. Rev.*, 1951, **58**, 332-347.
6. HASTORF, A. H., & CANTRIL, H. They saw a game: A case study. *J. Abn. & Soc. Psychol.*, 1954, **49**, 129-134.
7. ITTELSON, W. H. The Ames Demonstrations in Perception. Princeton: Princeton Univ. Press, 1952.
8. JACKSON, D. N. A further examination of the rôle of autism in a visual figure-ground relationship. *J. of Psychol.*, 1954, **38**, 339-357.
9. KILPATRICK, F. Human behavior from the transactional viewpoint. Hanover, N. H.: Inst. for Assoc. Research, 1952.
10. SCHAFER, R., & MURPHY, G. The rôle of autism in a visual figure-ground relationship. *J. Exper. Psychol.*, 1943, **32**, 335-343.
11. SMITH, D. E., & HOCHBERG, J. E. The effect of "punishment" (electric shock) on figure-ground perception. *J. of Psychol.*, 1954, **38**, 83-87.
12. WITTEICH, W. J. An investigation into the nature of aniseikonic distortion. Ph.D. dissertation, Princeton Univ., 1954.

The Menninger Foundation
Topeka, Kansas

THE EFFECT OF INCREASED TRIALS ON THE STABILITY OF THE PGR IN AN AVOIDANCE SITUATION*

S. NOVAK, D. MORIARTY, M. KOLTUV, AND L. WELCH

Department of Psychiatry, Cornell University Medical College, and the Payne Whitney Psychiatric Clinic, New York Hospital

A. INTRODUCTION

In a previous study (1), two groups were tested in a mock reaction time situation in which one group could avoid a threatened horn blast. The psychogalvanic responses were measured at various points in the procedure, the only significant differences being the greater mean PGRs of the Avoidance Group to the differential instructions.

The purpose of the present study is to investigate the effect of varying the number of trials preceding the differential instructions on the psychogalvanic reactivity of both groups.

B. SUBJECTS

The subjects included 87 college students, 77 females and 10 males, divided into six groups: Group Ia, the three-trial Avoidance group ($N = 15$); Group Ib, the three-trial Non-avoidance group ($N = 12$); Group IIa, the nine-trial Avoidance group ($N = 15$); Group IIb, the nine-trial Non-avoidance group ($N = 15$); Group IIIa, the 15-trial Avoidance group ($N = 15$); Group IIIb, the 15-trial Non-avoidance group ($N = 15$). Groups Ia and Ib were the original groups cited in the previous study (1).

C. APPARATUS

Mock reaction time apparatus consisted of a depressable microswitch lever attached to the armrest of a chair in which subjects were seated. A Klaxon automobile horn which was never sounded was conspicuously placed two feet to the left of the testing chair. The subject faced a ground glass screen at a distance of five feet. From the adjacent control room, the examiner controlled a green signal light (25 watts) and could also instruct the subject by means of an intercommunication system.

The instrument used to record the PGR was a Fordham two-stage DC amplifier in series with an Esterline-Angus graphic ammeter which provided

*Received in the Editorial Office on October 26, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

a permanent record. Thirty-two mm. bakelite cup electrodes, $\frac{3}{8}$ of an inch in height, filled with Cambridge electrode jelly, were used.

D. PROCEDURE

All of the subjects were prepared for testing in the same manner. Palms were cleansed with acetone and rubbed briskly for one minute with electrode jelly and a tongue depressor. Jelly was then wiped off and jelly-filled cups strapped to the palms. The subject remained in this "soak" condition for 10 minutes. Then the examiner demonstrated the operation of the reaction time lever, making certain that the subject noticed the horn. All further instructions were given over the loudspeaker from the control room. Initial instructions given to all subjects were as follows:

This is a simple reaction time test. The main purpose is to record your reaction time. In a few minutes you will see a green light flash behind the ground glass window in front of you. This signals the start of the test. About 10 seconds after this the light will re-appear. Press the lever with your index finger as soon as you see this second flash of light.

In Groups Ia and Ib, three trials were given after the initial instructions; Trials 2 and 3 were preceded by the examiner saying: "we will now repeat the same procedure once again." In Groups IIa and IIb, nine trials were given after the initial instructions, each of which was preceded by the same statement by the examiner. Groups IIIa and IIIb received 15 trials after the initial instructions which were also given in the same manner. After the appropriate number of preliminary trials, the critical or differential instructions were presented as follows:

To Groups Ia, IIa, and IIIa—The Avoidance Groups:

We are going to repeat the same procedure. Watch for the starting signal. However, from the variation in your reaction time we know you are able to respond faster than you are doing. If you don't press fast enough, the auto horn to your left will be sounded. You can escape hearing it by reacting at the required speed. Remember! React just as fast as you can.

To Groups Ib, IIb, and IIIb—The Non-Avoidance Groups:

We are going to repeat the same procedure. Watch for the starting signal. However, from the variation in your reaction time we know you are able to respond faster than you are doing. From now on the auto horn to you left will be sounded with the presentation of the second green light.

One further trial was given to each group following these instructions.

Of course, reaction time was never measured nor was the horn threat ever carried out.

E. RESULTS AND DISCUSSION

PGRs of all groups were compared and tested for the significance of any differences. Individual *PGRs* were computed in terms of the magnitude of the difference between each subject's resistance level prior to the critical instructions and the peak of his *PGR* after the presentation of these instructions. This difference was expressed in ohms and means for each group were obtained (Table 1). From the three-trial Avoidance group a mean of 902.00 ohms was obtained, and from the three-trial Non-avoidance group a mean of 273.33 ohms, the difference being significant at the .01

TABLE 1
MEAN *PGRs* TO THE DIFFERENTIAL INSTRUCTIONS

	I (3 trials)	II (9 trials)	III (15 trials)
The "a" Groups (Avoidance)	902.00 ohms	940.66 ohms	922.00 ohms
The "b" Groups (Non-Avoidance)	273.33 ohms	523.33 ohms	901.33 ohms
Difference	628.67 ohms	+17.33 ohms	20.67 ohms
Significance	.01	.025	—

level as previously reported (1). The nine-trial Avoidance group gave a mean *PGR* of 940.66 ohms and the nine-trial Non-avoidance group a mean of 523.33 ohms, the difference being significant at the .025 level. The 15-trial Avoidance group gave a mean *PGR* of 922.00 ohms and the 15-trial Non-avoidance group a mean of 901.33 ohms, the difference not being significant.

As is shown in Table 1, the psychogalvanic reactivity of the Avoidance groups remained constant in magnitude despite the number of trials used. This stability, we believe, may be explained in terms of the rationale given in the previous study; namely, the immediate autonomic alerting for the avoidance of trauma. The decreasing difference between the mean *PGRs* to the three-, nine-, and 15-trial differential instructions is a function of the increasing reactivity of the Non-avoidance groups, and may reflect attitudinal changes occurring as the experiment progresses. Subjects in the 15-trial groups, questioned after testing, reported feelings of diminished attention and extreme boredom. It might well be that these mental states served to increase significantly the reactivity of the Non-avoidance group as the number of trials was increased by progressively increasing the stimulus value of the Non-avoidance instructions.

F. SUMMARY AND CONCLUSIONS

Eighty-seven subjects were divided into six groups as follows: The three-trial Avoidance group ($N = 15$); the three-trial Non-avoidance group ($N = 12$); the nine-trial Avoidance group ($N = 15$); the nine-trial Non-avoidance group ($N = 15$); the 15-trial Avoidance group ($N = 15$); and the 15-trial Non-avoidance group ($N = 15$). After the appropriate number of trials, differential instructions were given informing the Avoidance groups that by quicker reactions an aversive auditory stimulus (Klaxon horn) could be avoided. The Non-avoidance groups were told that they would hear the horn on the next trial. Mean *PGRs* to the differential instructions were obtained for all groups.

The reactivity of the Avoidance groups remained stable; however, the reactivity of the Non-avoidance groups increased as the number of trials prior to the differential instructions increased, the difference thereby decreasing. At three trials the difference was significant at the .01 level. At nine trials the significance diminished to .025 and no significance was found between the 15-trial groups. It appears that the reactivity of subjects in the Non-avoidance situation is at least in part a function of the number of trials which precede the critical instructions.

REFERENCES

1. NOVAK, S., HAYES, R., GOODMAN, J., & WELCH, L. The effect of an avoidance situation on the psychogalvanic response. *J. of Psychol.*, 1955, 40, 307-311.

Payne Whitney Psychiatric Clinic
The New York Hospital
525 East 68th Street
New York City 21

EFFECTS OF ESTHETIC SURROUNDINGS: I. INITIAL EFFECTS OF THREE ESTHETIC CONDITIONS UPON PERCEIVING "ENERGY" AND "WELL-BEING" IN FACES*¹

Department of Psychology, Brandeis University

A. H. MASLOW AND N. L. MINTZ

A. THE PROBLEM

Esthetically sensitive individuals together with city planners, art educators, and related workers have long been intuitively aware of the effects of esthetic surroundings. Yet as far as we know there have been no experimental studies published on the effects of beautiful and ugly environments upon people. Surveys of the experimental esthetics (1, 3), color (12), and art (5) literature show research to be centered on "formal" properties of rhythm, style, color, line, etc., color preference and personality studies, color-concept matching experiments, and projective technique and art therapy work. We have found research on the effects of music (9, 16, 17) and color (2, 4, 6, 7, 10, 11, 13, 14, 15, 18) to be focused on the behavioral consequences of different melodic styles or hues *per se*, but not on music or color as part of the complex esthetic environment. The present experiment was undertaken as an initial step in studying the effects of beauty and ugliness upon people. It tested the short-term effects of three visual-esthetic conditions: "beautiful," "average," and "ugly" rooms.

B. METHOD

Three rooms were used. The "beautiful" room (BR) impressed people as "attractive," "pretty," "comfortable," "pleasant." It was 11' x 14' x 10' and had two large windows, beige-colored walls, an indirect overhead light, and furnishings to give the impression of an attractive, comfortable study. Furnishings included a soft armchair, a mahogany desk and chair combination, two straight-backed chairs, a small table, a wooden bookcase, a large Navajo rug, drapes for the windows, paintings on the walls, and some sculp-

*Received in the Editorial Office on November 4, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹This research was supported by Brandeis University. We wish to express our thanks to R. Held and R. B. Morant for their helpful discussion and assistance, and J. Glick for his photography work. Another paper will present the more complex data obtained from subjects moving into a second esthetic condition.

ture and art objects on the desk and table. These were all chosen to harmonize as pleasantly as possible with the beige walls.² The "ugly" room (*UR*) evoked comments of "horrible," "disgusting," "ugly," "repulsive." It was 7' x 12' x 10' and had two half-windows, battleship-gray walls, an overhead bulb with a dirty, torn, ill-fitting lampshade, and "furnishings" to give the impression of a janitor's storeroom in disheveled condition. There were two straight-backed chairs, a small table, tin cans for ashtrays, and dirty, torn window shades. Near the bare walls on three sides were such things as pails, brooms, mops, cardboard boxes, dirty-looking trash cans, a bed-spring and uncovered mattress, and assorted refuse. The room was neither swept nor dusted and the ashtrays were not emptied. The "average" room (*AR*) was a professor's office 15' x 17' x 10', with three windows, battleship-gray walls, and an indirect overhead light. Furnishings included two mahogany desk and chair combinations, two straight-backed chairs, a metal bookcase, window shades, a metal filing cabinet, and a cot with a pleasant-looking green bedspread. It gave the appearance of a clean, neat, "worked-in" office in no way outstanding enough to elicit any comments. To help restrict room differences to the visual mode, the experiment was done in the evening when the building was quiet; the *S*'s chair in the three rooms was of the same type; the rooms were well-lit (though *UR* had direct, harsh light); and the windows were always open, preventing the dust and dirt in *UR* from developing a musty odor.



FIGURE 1

SAMPLES OF THE NEGATIVE-PRINT PHOTOGRAPHS (Reduced 75 per cent)

A six-point, two dimension rating scale was used to test the effects of the conditions upon an *S*'s judgment of 10 negative-print photographs (see Figure 1). The dimensions to be rated were "energy" and "well-being" for each photograph. The rating scale thus had 10 judgments per dimension, each judgment with a weight of from 1-6. Summing the dimensions sepa-

²We wish to thank B. Maslow for her assistance in this.

rately would give two total scores, each having a possible range of from 10-60. These totals were averaged, giving an average "energy" and "well-being" score for each *S*. This average score could likewise range between 10 and 60. The 10 photographs were arranged alternately male and female, with two dummy extras³ preceding and following this series. Duplicate series were used for the three rooms.

As each *S* was met by the interviewer (NLM), he was told approximately the following:

We are conducting an experiment on facial stereotypy. You are familiar with Shakespeare's Cassius who had a lean and hungry look; this is an example of facial stereotypy. There cannot be any right or wrong answers as we are interested in the *impressions* faces give you.

At this point Köhler's expressive-line figures (8, p. 225) were demonstrated.

In just the same way as these lines appeared to have particular concept characteristics, we think faces will have certain trait characteristics. You are going to see negative prints like this sample. By negative printing, and dressing the people in this unusual fashion, we minimized hairline, clothing, and expression and emphasized bone structure and shape. We want you to give your impressions of these faces, similar to the way you gave impressions of the lines shown previously.

The *BR* and *UR* *Ss* were sent to their respective rooms to be tested by a naïve examiner who also thought the experiment was on facial stereotypy. The interviewer brought the *AR* *Ss* to their room and tested them. This elaborate prelude served two purposes. It insured the naïvete of the *Ss* and examiners (just one *S* guessed the purpose, and then only when thinking about the test a few days later), and it helped to reduce tension.

A test of visual-esthetic environment should emphasize spontaneity and informality or else task orientation or test anxiety may reduce the effects. This was demonstrated in a pilot study of similar design to this one,⁴ which failed to show significant differences between conditions in part due to the *Ss* anxiety and task orientation; they hardly looked away from the test material. Therefore, as each *S* entered the room the examiner was called out on some pretext. The *S* was left in the room for two minutes, allowing him to "soak" in the visual field. When the examiner returned he engaged the *S* in a rambling discussion of "fatigue energy" and "displeasure/well-being" with the intention of getting the *S* to name the moods

³Used for purposes to be discussed in another paper.

⁴Conducted by B. Maslow and A. H. Maslow.

just discussed. By allowing the *S* to choose his own concepts instead of being given our concepts for the dimensions, we felt there was greater likelihood of achieving a common semantic process among the *Ss*. Assuming the *S* chose "weary/zestful" and "irritable/content," the examiner then continued:

Now I would like you to tell me if this first face looks slightly, rather, or very weary, or slightly, rather, or very zestful. Then do the same for "irritable/content."

This was done for each of the 10 faces plus the four dummy extras. The *S* was encouraged to give any other impressions, and task-interrupting, idle conversation initiated by the examiner kept the atmosphere informal. The scores were marked on a scoresheet by the examiner. The *S* was in the room with the examiner at least 10 minutes additional to the time spent in the room alone.

Twenty-six male and 16 female undergraduate Brandeis University students volunteered for this experiment. They were recruited at large and simply told that we wanted them for "a study in faces and traits." Sixteen were for the *BR* group, 16 for the *UR* group, and 10 for the *AR* group. There were an even number of males and females in *BR* and *UR*, only males in *AR*. A naïve male and a naïve female were hired to examine the *BR* and *UR* *Ss*; one of the authors (NLM) examined the *AR* group.

The following controls were used with the *BR* and *UR* groups. Each experimenter tested eight *Ss* in the *BR* group and eight in the *UR* group, four of which were of one sex and four of the other. Half of the *Ss* in each group were asked to rate "energy" first and half to rate "well-being" first. After these groups were tested, the *AR* group was added to give additional information. Analysis of the *BR* and *UR* groups indicated these controls would be unnecessary,⁵ so all *AR* *Ss* were males, tested by a male examiner, and asked to rate "energy" first.

Our hypotheses were that scores obtained in the "beautiful" room would be higher (more "energy" and "well-being") than those in either the "average" or "ugly" rooms, and that scores in the "ugly" room would be lower (less "energy" and "well-being") than those in either the "average" or "beautiful" rooms.

⁵A four-way classification analysis of variance was done on *BR* and *UR* for rooms, experimenters, sexes, and rating order. The scores for rooms were significantly different, but those for experimenters, sexes, and rating order were not, nor did the four variables interact significantly.

C. RESULTS

Table 1 gives the results of an analysis of variance on the differences in the scores obtained in the three rooms. The scores for the three rooms were significantly different, as shown by the F ratio. Since the variances were not significantly heterogeneous, the F indicates a significant difference in the means. Table 2 shows these differences. The average ratings for "energy" and "well-being" in BR were significantly higher (beyond the .001 level) than ratings in UR , and significantly higher (beyond the .05 level) than ratings in AR . The average ratings in AR were higher, but not significantly so, than ratings in UR .

TABLE 1
DIFFERENCES IN SCORES OBTAINED FOR THREE ROOM CONDITIONS BASED ON AN
AVERAGE SCORE FOR EACH OF 42 S s

Source of variation	df	Variance estimate	F	p
Between rooms	2	153.25	6.49	<.01
Within rooms	39	23.63		
Total	41			

TABLE 2
DIFFERENCES BETWEEN THE MEANS OF SCORES IN THREE ROOM CONDITIONS

Room	Mean	Compared to room	Mean difference	t^*	p^{**}
UR	31.81	AR	2.19	1.12	<.30
AR	34.00	BR	3.99	2.04	<.05
BR	37.99	UR	6.18	3.54	<.001

*The within rooms variance estimate of Table 1 was used as the estimate of the standard error of the difference.

**Since the standard error was based on 39 df , the t was entered as a CR in the normal probability table.

Since an average score below 35 would indicate the S generally rated the 10 faces as "fatigued" and "displeased," while one above 35 would indicate the S rated the faces as having "energy" and "well-being,"⁶ Table 2 indicates a second result. It can be seen that the mean for the UR group is within the "fatigued" and "displeased" range; the mean for AR at the upper limit of the "fatigued" and "displeased" range; and the mean for the BR group within the "energy" and "well-being" range.

⁶ A ρ based on the 42 S s for the separate dimension totals ("fatigue/energy" and "displeasure/well-being") showed the dimensions to have a positive correlation of .79, significant beyond the .001 level.

D. DISCUSSION

We may summarize the results as follows. The *Ss* in our "beautiful" room gave significantly higher ratings (more "energy" and "well-being") than *Ss* in either the "average" or "ugly" rooms. Also, while the mean for the scores in the "beautiful" room fell in the "energy" and "well-being" range, the means for the other two groups fell in the "fatigued" and "displeased" range, indicating a qualitative difference in the group scores. We can be rather confident that the difference between the scores obtained in *BR* and *UR* is reliable. While the scores in *AR* are significantly lower than those in *BR* and somewhat higher than those in *UR* (results which are in the expected order), we cannot be as confident of where, between *BR* and *UR*, the *AR* group is placed. Recognizing the situational nature of our definitions of beauty, average and ugly, there still are interesting implications if our research would continue to find the effects of "average" surroundings to lie closer to those of "ugly" than to those of "beauty," rather than finding that effects of "average" lie midway between the two, or closer to "beauty." This, of course, would have immediate relevance for professors and their offices.

While many questions remain to be answered by research now in progress, certain points may be noted at the present time. We may begin by excluding the possibility that differences between groups resulted from suggestion or a "rôle-playing" attitude assumed by the examiners or the *Ss*. Indirect interviewing of the examiners after each day's testing, and each *S* after being tested, assured us that the examiners and *Ss* continued to be unaware of the experimental purpose. The controls for noise, odor, time of day, type of seating, examiners, etc., make us rather confident that the potent factor lay in the visual-esthetic qualities of the three rooms.

Regarding the effects obtained, a number of problems come to mind, some of which will be treated in later papers. Were these merely short-term effects; would the *Ss* adapt to the rooms with time and negate the initial differences obtained? How many individuals in each group were affected by the conditions? Were the *Ss* affected by the rooms *per se*? The possibility also exists that the results could have been obtained via the effect of conditions upon the examiners. This, of course, would not change the major implication of the findings; it would shift the emphasis from the rooms having a short-term effect directly upon the *Ss* to their having a long-term effect upon the examiners, which sufficiently affected the interpersonal relations between examiner and *S* so as to cause differences in group scores irrespective of which examiner was present.

In considering what may be the "potent" visual-esthetic aspects of the rooms, we may tentatively exclude as crucial in themselves the differences between room sizes, and neatness, orderliness, or cleanliness. Although *UR* was the smallest, *AR* was the largest; although *UR* was dirty and messy, *AR* was clean and neat. Both *UR* and *AR* had gray walls and cold colors in contrast to beige walls and warm colors in *BR*. While this may be important for understanding the difference between *BR* and *AR* scores, by itself it would not explain the possibly genuine difference between *AR* and *UR* scores. At present the most reasonable conclusion appears to be that all of these aspects were operating to produce three esthetically different-looking rooms, which in the case of "beautiful" and "ugly" resulted in clear differences between *Ss* ratings of the "energy" and "well-being" of faces.

E. SUMMARY

An experiment was conducted as an initial step in studying the effects of esthetic surroundings upon people. Three visual-esthetic conditions were used: "beautiful," "average," and "ugly" rooms. In each room, subjects unaware of the experimental purpose were asked to rate the "fatigue energy" and "displeasure/well-being" of 10 negative-print photographs of faces. The results were: (a) the group in the "beautiful" room gave significantly higher ratings (more "energy" and "well-being") than groups in either the "average" or "ugly" rooms; (b) the "average" room group had somewhat higher ratings than the "ugly" room group; (c) the mean score for ratings in the "beautiful" room fell in the "energy" and "well-being" range, while the means for the ratings in the other two rooms fell within the "fatigued" and "displeased" range. Discussion pointed out that: (a) suggestion, "role-playing," or variables other than visual-esthetic ones did not account for the differences obtained; (b) there seems at present to be no single visual-esthetic quality that can account for the differences among all three groups; (c) the effects may possibly have been obtained by the rooms' affecting the subject-examiner relationship.

REFERENCES

1. CHANDLER, A. R., & BARNHART, E. N. *A Bibliography of Psychological and Experimental Esthetics, 1864-1937*. Berkeley: Univ. California Press, 1938.
2. DEUTSCH, F. Psycho-physical reactions of the vascular system to influences of light and to impressions gained through light. *Folia Clinica Orientalia*, 1937, 1, Faces. 3 & 4.
3. DROUGHT, R. A. A survey of studies in experimental esthetics. *J. Educ. Res.*, 1929, 20, 97-102.
4. EHRENWALD, N. Referred to by Ellinger, F. *The Biologic Fundamentals of Radiation Therapy*. New York: Elsevier, 1941.

5. FAULKNER, R. A survey of recent research in art and art education. In G. M. Whipple (Ed.) *Art in American Life and Education, Fortieth Year-book, NSSE*. Chicago: Univ. Chicago, 1941. (Pp. 369-377.)
6. GOLDSTEIN, K. Some experimental observations concerning the influence of colors on the function of the organism. *Occup. Ther.*, 1942, **21**, 147-151.
7. GOLDSTEIN, K., & ROSENTHAL, O. Zum Problem der Wirkung der Farben auf den Organismus. *Schweiz. Arch. Neurol. & Psychiat.*, 1930, **26**, 3-26.
8. KÖHLER, W. *Gestalt Psychology*. New York: Liveright, 1929.
9. MITCHELL, S. D., & ZANKER, A. The use of music in group therapy. *J. Ment. Sci.*, 1948, **94**, 737-738.
10. MITRA, S. C., & DATTA, A. The influence of color on the estimation of area. *Indian J. Psychol.*, 1939, **14**, 91-94.
11. MOGENSEN, M. F., & ENGLISH, H. B. The apparent warmth of colors. *Amer. J. Psychol.*, 1926, **37**, 427-428.
12. NORMAN, R. D., & SCOTT, W. A. Color and affect: A review and semantic evaluation. *J. Gen. Psychol.*, 1952, **46**, 185-223.
13. PIERCE, D. H., & WEINLAND, J. D. The effect of color on workmen. *Person. J.*, 1934, **13**, 34-38.
14. PRESCOTT, D. B. Psychological analysis of light and color. *Occup. Ther.*, 1942, **21**, 135-146.
15. PRESSEY, S. L. The influence of color upon mental and motor efficiency. *Amer. J. Psychol.*, 1921, **32**, 326-356.
16. RUBIN, H. E., & KATZ, E. Aurotone films for the treatment of psychotic depressions in an army general hospital. *J. Clin. Psychol.*, 1946, **2**, 333-340.
17. SCHULLIEN, D. M., & SCHOEN, M. *Music and Medicine*. New York: Schuman, 1948.
18. WALLIS, W. A. The influence of color on apparent size. *J. Gen. Psychol.*, 1935, **13**, 193-199.

Department of Psychology
Brandeis University
Waltham, Massachusetts

THE VARIABILITY OF REACTION TIME IN THE SENSORI-MOTOR INDUCTION SYNDROME WITH SPECIAL REFERENCE TO THE EFFECT OF COLORS*

*Department of Nervous Diseases of the Rothschild Hadassah-University Hospital,
Hebrew University-Hadassah Medical School, Jerusalem, Israel*

L. HALPERN AND S. KUGELMASS

A. THE PROBLEM

The sensorimotor induction syndrome in unilateral disequilibrium is based on the evidence that changes in postural pattern may induce changes in sensory functions (1-3, 4-5, 12). These perceptual changes are phenomenologically and physiologically of a special nature occurring in spite of the anatomical integrity of the sense organs. These changes appear in the sense organs of that side of the body which corresponds to the side of the disequilibrium. The resulting sensory changes manifest themselves as primary symptoms of perceptual displacement in visual, tactile, and haptic spheres, as well as secondary symptoms in the form of qualitative and quantitative alterations of perception (6). It was further shown that this syndrome is characterized by the mutual influence of sensory and motor functions (7-10), and that the given features are affected differently by the opening and closing of both eyes and by the alternating closure of each eye (11). It was finally found that the various motor and sensory disturbances are influenced by colors. In connection with these results it seemed of interest to study the influence of the same conditions on reaction time in these patients.

B. SUBJECTS AND METHOD

Four patients manifesting the symptoms of sensorimotor induction syndrome in unilateral disequilibrium served as subjects. All of these patients had been studied carefully previously, and were accustomed to experimental sessions. The syndrome appeared on the left side of the body in three of the patients and on the right side in the fourth.

The patient was seated in a comfortable chair to the arm of which was attached a microswitch. This microswitch was connected to the clutch of a Standard clock which measured time in units of .01 of a second. The ex-

*Received in the Editorial Office on November 15, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

perimeter would, after giving a warning signal, throw a switch which would present either a light or sound stimulus. This switch which powered the light or buzzer was connected to the recording clock as well, and would simultaneously present the stimulus and start the clock. The clock would stop when the subject pressed the microswitch. Thus, the reading on the clock was a measure of the time that elapsed between the presentation of the stimulus and of the patient's response, that of pressing on the microswitch. Only one stimulus, either the light or the sound, was used consistently during an experimental session. An experimental session lasted about one hour, during which measures of reaction time were obtained under the following conditions: (a) both eyes open, (b) both eyes closed, (c) alternating closure of each eye, and (d) with red, blue, and in some cases yellow filters placed in front of the open eye. A rough calibration of the filters, by means of a spectrophotometer, indicates that the bulk of the transmission for the blue filter is from 4,000-5,000A°, while for the red filter it is from 6,000-7,000A°. The yellow filter's transmission lies between these two, at 5,000-7,000A°. A gross comparison of the filters in terms of luminance or brightness indicates that the blue filter has the greatest density, the yellow the least density, with the red intermediate. The patient wore a set of frames over her eyes into which the colored filters could be placed. In order to ensure the "closing" of an eye, a black cloth was tied comfortably over it.

A series of 10 trials was undertaken under each experimental condition. The only exception to this was when a patient, under the influence of red, became so uncomfortable that the series had to be cut short. Rest periods were given to the patients between conditions to minimize fatigue effects.

Certain additional series of reaction time measurements were recorded. The first part of these series included measurements with a red filter placed in front of the open homolateral eye, which corresponds to the side of the disequilibrium. After about 10 or so reaction time trials, the red filter was quickly replaced by a blue filter and further reaction time trials were recorded. The data from these series are presented in graphs.

C. RESULTS

The means and standard deviations of all of the series were calculated. All but two of the series include 10 trials. Values of "*t*" were computed to test the significance of differences between the means (averages) as well as significance of differences between the standard deviations (measure of variability) of the various series. Differences are reported as significant if they reach the .05 level of confidence.

1. Individual Results

a. *Patient D.A. with sensorimotor induction syndrome in left-sided disequilibrium.* A comparison of the various series for Patient D.A. reveals the following significant differences: Light stimulus—the mean of the series during which both eyes were closed (38.80) is significantly greater than the mean of the series during which both eyes were open (30.70). All

TABLE 1

	Both Eyes Open	Both Eyes Closed	Contra- lateral Eye Open	Contra- lateral Eye Blue	Contra- lateral Eye Red	Contra- lateral Eye Yellow	Homo- lat. Eye Open	Homo- lat. Eye Blue	Homo- lat. Eye Red
<i>Light stimulus</i>									
Mean	30.70	38.80	24.40	23.50	22.70	24.10	31.70	37.70	36.10
St. Dev.	2.76	3.22	1.96	2.66	2.41	2.38	5.81	4.88	4.77
<i>Sound stimulus</i>									
Mean	17.70	17.00	16.40	17.20	16.30	16.30	35.80	21.70	21.80
St. Dev.	1.49	2.53	1.86	1.33	2.56	1.55	14.34	2.72	2.23

of the means of the series during which the homolateral left eye which corresponds to the side of the disequilibrium was open are significantly greater than the means of the series during which the right contralateral eye was open. Thus the means of the series during which the left homolateral eye was open (31.70), open and covered with a blue filter (37.70), and open and covered with a red filter (36.10), are all significantly greater than the means of the series during which the right contralateral eye was open (24.40), open and covered with a blue filter (23.50), open and covered with a red filter (22.70), or open and covered with a yellow filter (24.10). The mean of the series during which the left homolateral eye was covered with a blue filter (37.70) is significantly greater than the mean of the series during which the left homolateral eye was open and uncovered (31.70).

The standard deviation of the series during which the left homolateral eye was open and uncovered (5.81) is significantly greater than the standard deviation of the series during which the right contralateral eye was open and uncovered (1.96), open and covered with a blue filter (2.66), open and covered with a red filter (2.41), and open and covered with a yellow filter (2.38). Also, the standard deviation of the series during which the left homolateral eye was open and uncovered (5.81), open and covered by a blue filter (4.88), and open and covered by a red filter (4.77), are all significantly greater than the standard deviation of the series during which the right contralateral eye was open and uncovered (1.96).

Sound stimulus—all of the means of the series during which the left homolateral eye was open are significantly greater than all of the means of the series during which the right contralateral eye was open. Thus, the means of the series during which the left homolateral eye was open and uncovered (35.80), was open and covered by a blue filter (21.70), and was open and covered by a red filter (21.80), are significantly greater than the means of the series during which the right contralateral eye was open and uncovered (16.40), open and covered by a blue filter (17.20), open and covered by a red filter (16.80), and open and covered by a yellow filter (16.30). The mean of the series during which the left homolateral eye was open and uncovered (35.80) is significantly greater than the means of the series during which the left homolateral eye was open and covered by a blue filter (21.70), or open and covered with a red filter (21.80).

The standard deviation of the series during which the left homolateral eye was open and uncovered (14.34) is significantly greater than the standard deviations of all of the other series undertaken using the sound stimulus.

No apparent difference between the reaction time values obtained with a red filter placed in front of the left homolateral eye, and those obtained immediately afterward with a blue filter placed in front of the same open eye were found in this patient.

b. Patient S.S. with sensorimotor induction syndrome in right-sided disequilibrium. A comparison of the various series for Patient S.S. reveals the following significant differences: The means of the series during which the right homolateral eye corresponding to the side of the disequilibrium, was open and uncovered (18.1), and open and covered by a red filter (21.3), are significantly greater than the mean of the series during which the right homo-

TABLE 2
LIGHT STIMULUS

	Both Eyes Open	Both Eyes Closed	Contra- lateral Eye Open	Contra- lateral Eye Blue	Contra- lateral Eye Red	Homo- lat. Eye Open	Homo- lat. Eye Blue	Homo- lat. Eye Red
Mean	16.9	21.2	14.7	15.0	20.6	18.1	15.7	21.3
St. Dev.	1.87	2.93	2.12	1.27	2.46	2.74	1.08	6.45

lateral eye was open and covered by a blue filter (15.7). The mean of the series during which the left contralateral eye was open and covered with a red filter (20.6) is significantly greater than the means of the series during which the left contralateral eye was open and uncovered (14.7), and open and covered with a blue filter (15.0). The mean of the series during which

the left contralateral eye was open and covered by a red filter (20.6) is significantly greater than the mean of the series during which the right homolateral eye was open and covered by a blue filter (15.7). The means of the series during which both eyes were open (16.9), and the right homolateral eye was open (18.1), are significantly greater than the mean of the series during which the left contralateral eye was open (14.7). The mean of the series during which both eyes were closed (21.2) is significantly greater than the mean of the series during which both eyes were open (16.9).

The standard deviations of the series during which the right homolateral eye was open and uncovered (2.74), and open and covered by a red filter (6.45), are significantly greater than the standard deviation of the series during which the right homolateral eye was open and covered by a blue filter (1.08). The standard deviation of the series during which the right homolateral eye was open and covered by a red filter (6.45) is significantly greater than the standard deviation of the series during which the left contralateral eye was open and uncovered (2.12), open and covered by a blue filter (1.27), or open and covered by a red filter (2.46), as well as the series during which the right homolateral eye was open and uncovered (2.47), and open and covered with a blue filter (1.08).

Figure 1 indicates that the replacement of a red filter by a blue filter in front of the right homolateral eye caused a clear progressive reduction of variability in the series of reaction time trials.

c. *Patient L.A. with sensorimotor induction syndrome in left-sided disequilibrium.* A comparison of the various series for Patient L.A. reveals the following significant differences: The mean of the series during which the left homolateral eye, corresponding to the side of the disequilibrium, was open and uncovered (83.8), and the mean of the series during which the

TABLE 3

	Both Eyes Open	Both Eyes Closed	Contra- lateral Eye Open	Contra- lateral Eye Blue	Contra- lateral Eye Red	Homo- lat. Eye Yel.	Homo- lat. Eye Open	Homo- lat. Eye Blue	Homo- lat. Eye Red
<i>Light stimulus</i>									
Mean	46.2		66.5	41.6	115.6		83.8	39.8	168.3
St. Dev.	14.77		14.29	6.86	54.60		19.63	7.96	112.4
<i>Sound stimulus</i>									
Mean	52.40	53.10	42.5	41.7	75.6	58.8	49.4	44.4	89.8
St. Dev.	10.41	21.64	12.24	11.08	19.01	21.77	20.89	15.49	35.88

right contralateral eye was open and uncovered (66.5), are both significantly greater than the mean of the series during which both eyes were open and

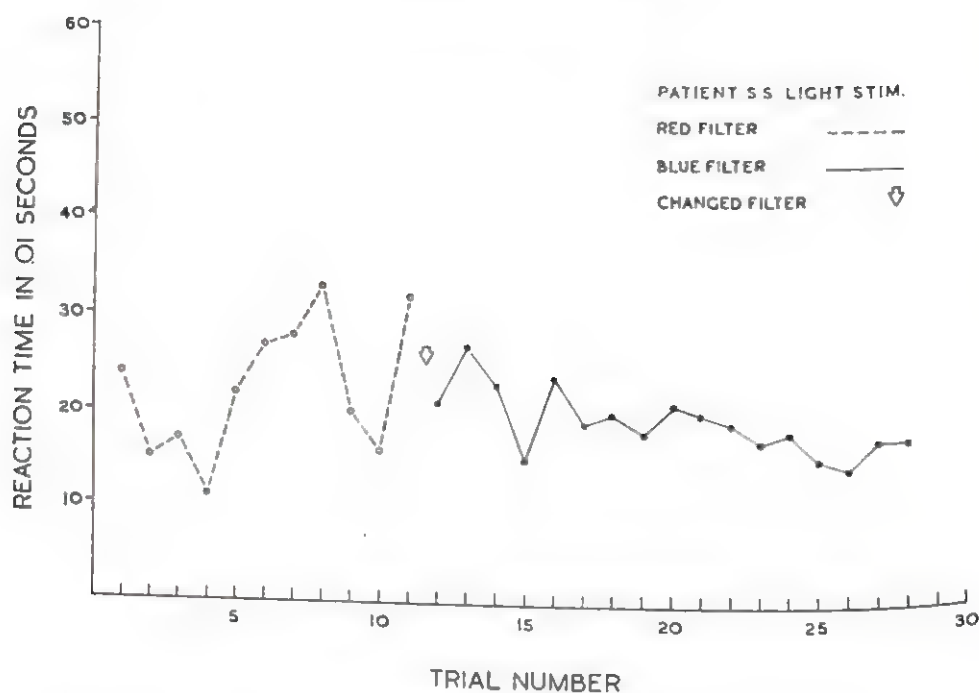


FIGURE 1
THE EFFECT OF REPLACING THE RED FILTER BY THE BLUE FILTER ON THE SPEED
OF REACTION TIME FOR PATIENT S.S.

uncovered (46.2). The mean of the series during which the left homolateral eye was open and uncovered (83.8) is significantly greater than the mean of the series during which the right contralateral eye was open and uncovered (66.5). The means of the series during which the right contralateral eye was open and covered by a red filter (115.6), and open and uncovered (66.5), are both significantly greater than the mean of the series during which the right contralateral eye was open and covered by a blue filter (41.6). The mean of the series during which the right contralateral eye was open and covered by a red filter (115.6) is significantly greater than the mean of the series during which the right contralateral eye was open and uncovered (66.5). The means of the series during which the left homolateral eye was open and covered by a red filter (168.3), and open and uncovered (83.8), are both significantly greater than the mean of the series during which the left homolateral eye was open and covered by a blue filter (39.8). The mean of the series during which the left homolateral eye was open and covered by a red filter (168.3) is significantly greater than the mean of the series during which the left homolateral eye was open and un-

covered (83.8). The means of the series during which the right contralateral eye was open and covered by a red filter (115.6), and open and uncovered (66.5), are both greater than the mean of the series during which the left homolateral eye was open and covered by a blue filter (39.8).

The standard deviation of the series during which the right contralateral eye was open and covered by a red filter (54.60) is significantly greater than both of the standard deviations of the series during which the right contralateral eye was open and uncovered (14.29), and open and covered by a blue filter (6.86). The standard deviations of the series during which the left homolateral eye was open and covered by a red filter (112.4), and open and uncovered (19.63), are both significantly greater than the standard deviation of the series during which the left homolateral eye was open and covered by a blue filter (7.96).

Figure 2 indicates that the replacement of the red filter by a blue filter in front of the left homolateral eye was followed by lesser reaction time values as well as a clear progressive reduction of variability. In this part of the experiment only three reaction time trials were taken with a red filter because of the discomfort to the patient.

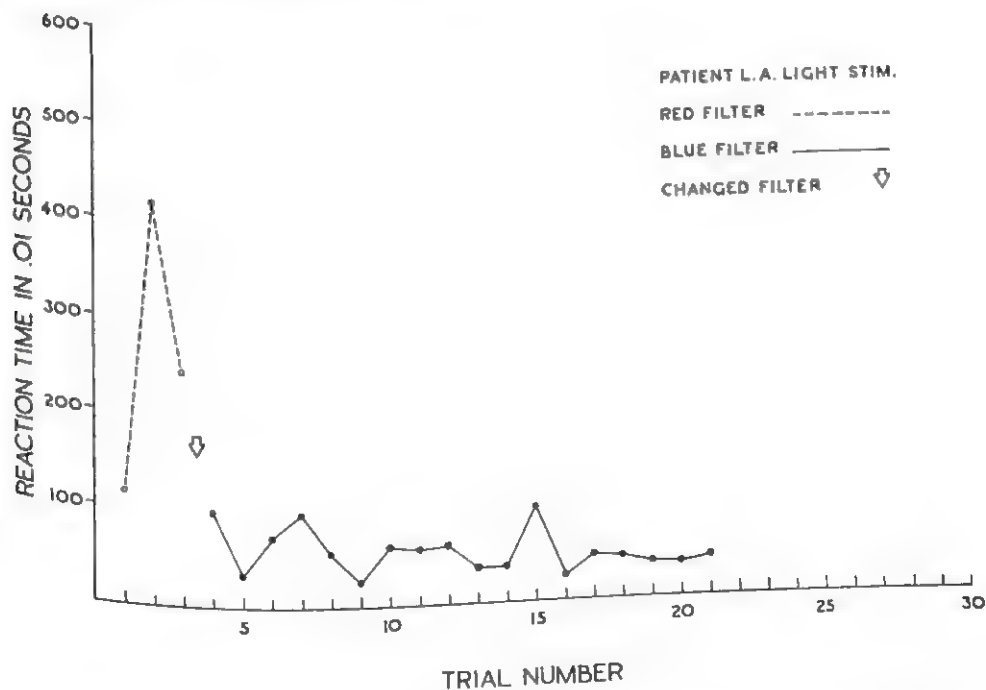


FIGURE 2

THE EFFECT OF REPLACING THE RED FILTER BY THE BLUE FILTER ON THE SPEED OF REACTION TIME FOR PATIENT L.A.

Sound stimulus—the mean of the series during which the left eye which corresponds to the side of the disequilibrium, was open and covered by a red filter (89.80) is significantly greater than the means of the series during which the left eye was open and uncovered (49.40) and open and covered by a blue filter (44.40). The mean of the series during which the right contralateral eye was open and covered by a red filter (75.60) is significantly greater than the means of the series during which the right contralateral eye was open and uncovered (42.5) and open and covered by a blue filter (41.70). The mean of the series during which the right contralateral eye was open and covered by a yellow filter (58.80) is significantly greater than the mean of the series during which the right contralateral eye was open and covered by a blue filter (41.70).

No significant differences were found in comparison of the standard deviations of the different series of trials.¹

d. *Patient A.F. with sensorimotor induction syndrome in left-sided disequilibrium.* A comparison of the various series for Patient A.F. reveals the following significant differences: Light stimulus—the mean of the series during which the left eye, corresponding to the side of the disequilibrium, was open and covered by a red filter (70.90) is significantly greater than the means of the series during which the left homolateral eye was open and

TABLE 4

	Both Eyes Open	Both Eyes Closed	Contra- lateral Eye Open	Contra- lateral Eye Blue	Contra- lateral Eye Red	Contra- lateral Eye Yellow	Homo- lat. Eye Open	Homo- lat. Eye Blue	Homo- lat. Eye Red
<i>Light stimulus</i>									
Mean	27.70		28.0	23.8	62.5	31.1	37.1	29.8	70.9
St. Dev.	7.76		8.66	3.68	17.03	7.91	7.80	8.86	28.9
<i>Sound stimulus</i>									
Mean	16.4	57.7	16.8	16.9	35.0	20.6	33.9	19.3	32.2
St. Dev.	4.10	14.33	1.94	5.17	11.52	5.26	9.83	4.27	8.68

uncovered (37.10), and open and covered by a blue filter (29.80). The mean of the series during which the right contralateral eye was open and covered by a red filter (62.50) is significantly greater than the means of the series during which the right contralateral eye is open and uncovered (28.00) and open and covered by a blue filter (23.80), and open and covered by a

¹With Patient L.A. during the series in which the homolateral left eye was open and covered by a red filter, there were seven trials with the light stimulus and five trials with the sound stimulus.

yellow filter (31.10). The mean of the series during which the left homolateral eye was open and uncovered (37.10) is significantly greater than the means of the series during which both eyes were open and uncovered (27.70) and during which the right contralateral eye was open and uncovered (28.00).

The standard deviations of the series during which the right contralateral eye was open and uncovered (8.66), and open and covered by a red filter (17.03), are both significantly greater than the standard deviation of the series during which the right contralateral eye was open and covered by a blue filter (3.68). The standard deviation of the series during which the left homolateral eye was open and covered by a red filter (28.90) is significantly greater than the standard deviations of the series during which the left homolateral eye was open and uncovered (7.80) or open and covered by a blue filter (8.86).

Figure 3 indicates that the replacement of the red filter by a blue filter in front of the open left homolateral eye was followed by shorter reaction time values, as well as a clear reduction of variability.

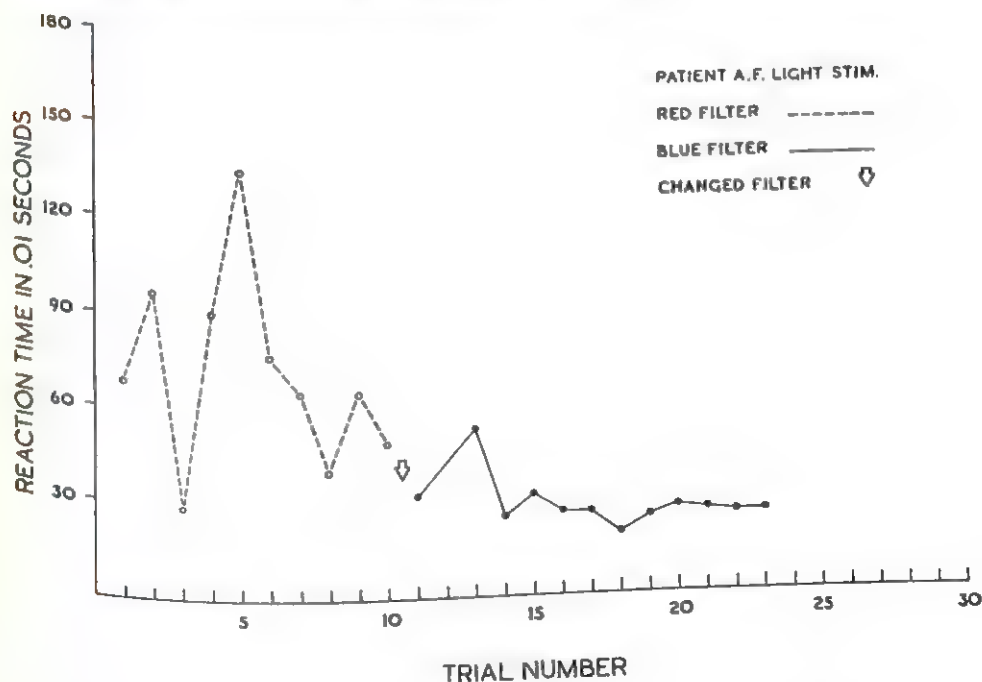


FIGURE 3
THE EFFECT OF REPLACING THE RED FILTER BY THE BLUE FILTER ON THE SPEED OF REACTION TIME FOR PATIENT A.F.

Sound stimulus—the means of the series during which the left homolateral eye was open and uncovered (33.90) and open and covered with a red filter (32.20) are significantly greater than the mean of the series during which the left homolateral eye was open and covered by a blue filter (19.30). The mean of the series during which the right contralateral eye was open and covered with the red filter (35.00) is significantly greater than the means of the series during which the right contralateral eye was open and uncovered (16.80), open and covered by the blue filter (16.90), and open and covered by the yellow filter (20.6). The mean of the series during which the right contralateral eye was open and covered by the yellow filter (20.6) is significantly greater than the mean of the series during which the right contralateral eye was open and uncovered (16.80). The means of the series during which the left homolateral eye was open and uncovered (33.90) and during which both eyes were closed (57.7) are both significantly greater than the means of the series during which both eyes were open (16.40) and during which the right contralateral eye was open and uncovered (16.80).

The standard deviation of the series during which the left homolateral eye was open and uncovered (9.83) is significantly greater than the standard deviation of the series during which the left homolateral eye was open and covered by the blue filter (4.27). The standard deviations of the series during which the right contralateral eye was open and covered with a blue filter (5.17), yellow filter (5.26), and the red filter (11.52), are all significantly greater than the standard deviation of the series during which the right contralateral eye was open and uncovered (1.94). The standard deviation of the series during which the right contralateral eye was open and covered by the red filter (11.52) is significantly greater than the standard deviations of the series during which the right contralateral eye was open and covered by the yellow filter (5.26) and the blue filter (5.17). The standard deviation of the series during which the left homolateral eye was open and uncovered (9.83) is significantly greater than the standard deviations of the series during which the right contralateral eye was open and uncovered (1.94) and of the series during which both eyes were open and uncovered (4.10). The standard deviation of the series during which both eyes were closed (14.33) is significantly greater than the standard deviation of the series during which both eyes were open and uncovered (4.10), and the series during which the right contralateral eye was open and uncovered (1.94).

2. Group Results

In addition to the individual analysis of data made above, it is of interest to consider the group trends:

a. Five comparisons of reaction times obtained with both eyes open may be made with reaction times obtained when both eyes were closed. Four of these five indicate greater mean values of reaction time under the experimental condition during which both eyes were closed. Three of these four differences are statistically significant. Five out of five of these comparisons indicate greater standard deviations during the condition in which both eyes were closed. One of these five is statistically significant.

b. Seven comparisons of reaction times obtained with the homolateral eye alone open may be made with reaction times obtained when the contralateral eye alone was open. All seven indicate greater mean values of reaction time under the condition when the homolateral eye alone was open. Six of these seven are statistically significant. Six of the seven comparisons indicate greater standard deviations during the condition in which only the homolateral eye is open. Three of these six are statistically significant.

c. Seven comparisons of reaction times obtained with only the contralateral eye open and uncovered may be made with reaction times obtained when the contralateral eye was open and covered with the blue filter. Four of these comparisons indicate greater mean values of reaction time under the condition during which only the contralateral eye was open and uncovered. One of these four was statistically significant. Five of the seven comparisons indicate greater standard deviations during the condition that only the contralateral eye was open and uncovered. Two of these five are statistically significant.

d. Seven comparisons of reaction times obtained with only the contralateral eye open and uncovered may be made with reaction times obtained when the contralateral eye was open and covered by the red filter. Five of these comparisons indicate greater mean reaction time values under the condition during which only the contralateral eye was open and covered by the red filter. All five are statistically significant. All seven of the comparisons indicate greater standard deviations during the condition that only the contralateral eye was open and covered by the red filter. Two of these seven are statistically significant.

e. Seven comparisons of reaction times obtained with only the homolateral eye open and uncovered may be made with reaction times obtained when only the homolateral eye was open and covered with the blue filter. Six of these seven comparisons indicate greater mean reaction times under the

condition during which the homolateral eye was open and uncovered. Four of these six were statistically significant. The seventh comparison which indicated greater mean reaction time during the condition when the homolateral eye was open and covered by the blue filter was significant. Six out of the seven comparisons indicate greater standard deviations during the condition that the homolateral eye was open and uncovered. Four of these six are statistically significant.

f. Seven comparisons of reaction times obtained with the homolateral eye open and uncovered may be made with reaction times obtained with the homolateral eye open and covered by the red filter. Six of these seven indicate greater mean reaction times under the condition during which the homolateral eye was open and covered by the red filter. Five of these six are statistically significant. The seventh comparison indicated a greater mean reaction time during the condition when the homolateral eye was open and covered by the red filter. Three of these four are statistically significant. Of the remaining three comparisons indicating greater standard deviations when the homolateral eye was open and uncovered, one was statistically significant.

g. Five comparisons of reaction times obtained with the contralateral eye open and covered by the blue filter may be made with reaction times obtained with the contralateral eye open and covered by the yellow filter. Three of these five comparisons indicate greater mean values of reaction time under the condition during which the contralateral eye was open and covered by the yellow filter. Two of these three are statistically significant. Four of the five comparisons indicate greater standard deviations during the condition when the contralateral eye was open and covered by the yellow filter.

h. Five comparisons of reaction times obtained with the contralateral eye open and covered by the red filter may be made with reaction times obtained with the contralateral eye open and covered by the yellow filter. Four of these five comparisons indicated greater mean reaction times with the contralateral eye open and covered by the red filter. Two of these four are statistically significant. Four of the five comparisons indicate greater standard deviations during the condition when the contralateral eye was open and covered by the red filter. Two of these four are statistically significant.

3. *Summary of Results*

Reaction time values are longer and more variable when both eyes are closed as compared with those obtained when both eyes are open. Reaction to the light stimulus with closed eyes in this experiment was possible because of the relative brightness used. In all of the trials reported the

patient stated that she was able to note the onset of the light stimulus clearly with both eyes closed. It is recognized that a contaminating factor is thus introduced since the intensity of the light stimulus is less with closed eyes than it is with open eyes. Control trials using two normals did not give different results with closed eyes than with open eyes under our experimental conditions.

Reaction time values are longer and more variable when the patient's homolateral eye is open as compared with those obtained when the contralateral eye is open.

Reaction time values are longer and more variable when a red filter is placed in front of the homolateral eye or even the contralateral eye. A blue filter, on the other hand, produces shorter and less variable reaction time values. A yellow filter produces results somewhat in between those obtained with the red and the blue filter. One of the patients, D.A., did not show any systematic sensitivity to different color filters.

D. DISCUSSION

The examination of reaction time in this study was carried out as mentioned on the basis of findings characteristic of the sensorimotor induction syndrome. It had been demonstrated in patients with this syndrome that the postural deviations of the whole body and the homolateral extremities in particular increased on closing of both eyes as compared with the original posture with eyes open. The closing of both eyes produced aggravation not only in the motor sphere, but in the sensory sphere as well. It was further shown that the existing disturbances were influenced differently by single closure of each eye. When the homolateral eye, corresponding to the side of the disequilibrium, was kept open, aggravation was produced similar to that obtained with both eyes closed. Keeping the contralateral eye open, on the other hand, induced an improvement and correction of the sensory and motor disturbance. In addition, a systematic influence of colors on the existing disturbances could be determined in these patients. While red affected a definite aggravation of the motor and sensory disturbances, they were corrected by blue. In its corrective effect green was close to blue, while yellow was closer to red in its deteriorating effect. In view of these findings, it appeared to be of interest to examine reaction time in these patients under the same conditions. This would provide a measure of the behavior of the whole organism with data that could be analysed graphically and statistically. The above mentioned results indicated in fact that, aside from the expected individual variations, clear differences in reaction time were noted under

the different experimental conditions. The results corresponded fundamentally with the earlier basic findings in which aggravation of both motor and sensory disturbances was noted when both eyes were closed or only the homolateral eye was open. Similarly, in this study, reaction time was generally slower, more variable, or both under these conditions.

In connection with the effect of colors, the present results also parallel the earlier findings. Reaction times to both visual and auditory stimuli were markedly longer and more variable while looking through the red filter. Reaction times were shorter and less variable when looking through the blue filter. The yellow filter produced an intermediate effect. These effects were more striking when the subject was looking through the homolateral eye, but they also appeared when the contralateral eye was being used. The question may be raised whether this "color" effect is based on the wave length of the light, the accompanying luminance or brightness, or some combination of both. While a definite answer will have to await further experimentation, our limited data point to the greater importance of the wave length. The extent of the slowing of the reaction time increased systematically from blue to yellow to red, or from the shorter to the longer wave length. The yellow filter having the least density or allowing the greatest luminance or brightness was associated with an intermediate effect. Thus the present results lie on a continuum that seems to be related to hue or spectrum rather than to increasing brightness.

Concerning the color effect, our results on reaction time correspond basically to the analogous effect of colors on the single functions. These results demonstrate that our general assumption about the "excitatory" effect of red and the "soothing" effect of blue is real and biologically founded.

E. SUMMARY

Four patients presenting the sensorimotor induction syndrome in unilateral disequilibrium were studied. Series of reaction time values were obtained under several experimental conditions. With all four patients reaction time values were longer and/or more variable when both eyes were closed as compared with both eyes being open. Reaction times were longer and/or more variable when the patient's homolateral eye was open as compared with those obtained when the contralateral eye was open. In three patients consistent differences were noted depending on the color of the filter. Red produced longer and/or more variable reaction time values, while blue produced shorter and/or less variable reaction time values. A yellow filter produced results somewhat intermediate. These results are the expression

of the neurophysiological mechanisms which underlie the sensorimotor induction syndrome.²

REFERENCES

1. HALPERN, L. The syndrome of sensorimotor induction in disturbed equilibrium. *Arch. Neurol. & Psychiat.*, 1949, **62**, 330.
2. ———. The syndrome of sensorimotor induction in combined cerebellar and labyrinthine injury. *J. Nerv. & Ment. Dis.*, 1951, **114**, 137.
3. ———. Le syndrome d'induction sensorimotrice dans les troubles de l'équilibre. Paris: Masson et Cie, 1951.
4. HALPERN, L., & KIDRON, D. F. Sensorimotor induction syndrome in unilateral disequilibrium. *Neurology*, 1954, **4**, 233.
5. HALPERN, L. Nouvelle contribution à l'étude du syndrome d'induction sensorimotrice dans la déséquilibration unilatérale. *Rev. Mensuelle de Psychiat. et de Neurol.*, 1955, **130**, 85.
6. ———. Secondary disturbances of perception. *J. Nerv. & Ment. Dis.*, 1952, **116**, 783.
7. HALPERN, L., & LANDAU, J. Head posture and visual functions. *Month. Rev. Psychiat. & Neurol.*, 1953, **125**, 148.
8. HALPERN, L. Head posture and sensory perception. Communication at the International Neurological Congress, Lisbon, 1953.
9. HALPERN, L., & FEINMESSER, M. Head posture and acoustic perception. *Month. Rev. Psychiat. et Neurol.*, 1954, **127**, 122.
10. HALPERN, L. Biological significance of head posture in unilateral disequilibrium. *Arch. Neurol. et Psychiat.*, 1954, **72**, 160.
11. ———. Optic function and postural attitude. *Neurology*, 1954, **4**, 830.
12. ROZANSKI, J. Contribution à l'étude du syndrome d'induction sensorimotrice. *Arch. Suisses Neurol. et Psychiat.*, 1952, **69**, 248.

Department of Neurology
Rothschild Hadassah University Hospital
Hebrew University
Jerusalem, Israel

²From the Department of Nervous Diseases of the Rothschild Hadassah-University Hospital, Psychophysiological Laboratory, Hebrew University-Hadassah Medical School. Supported by a grant of the Hebrew University.

A TACTUAL-KINESTHETIC CURVATURE ILLUSION*

Department of Psychology, Michigan State University

EUGENE S. EDGINGTON

It appears that in some situations the primary cue utilized by the blind in judging the curvature of a surface might be the movement of the blind person's hand toward and away from his body as he moves his hand across a vertical surface in front of him. If a surface could be moved toward and away from a blind or blindfolded observer, during the exploration with the hand, without the observer's knowledge, then the observer could be deceived in his judgment of the curvature of the surface. The following discussion will indicate how this type of movement of a surface can be produced without the observer's knowledge. The illusion produced is a tactual-kinesthetic curvature illusion in which a flat, smooth surface feels curved. The perceived radius of curvature and direction of curvature depend on the manner of moving the surface toward and away from the observer.

No apparatus is needed to produce this illusion. A blindfolded observer is told that he is to judge the shape of a surface by moving his finger across it. The experimenter (*E*) holds a flat smooth-surfaced object vertically between himself and the observer (*O*), and places *O*'s finger on the surface. As *O* continuously moves his finger in one direction across the surface, *E* moves the surface toward and away from *O*. *O* will feel the surface to have the curvature of the curve generated by his finger movement; e.g., if the surface moves toward, then away from *O*, he feels a convex surface. Since the perceived radius of curvature and direction of curvature are determined by the curve generated by the finger movement, *E* can alter the illusion by varying the movement of the surface toward and away from the observer, thus guiding the observer's finger in generating the curve. A moderately slow, continuous hand movement by *O* allows *E* to coordinate the movement of the surface with the hand movement. The best speed for *O*'s hand movement depends on the individual experimenter's preference, but he should indicate to *O* a desired rate of hand movement. *E* should not move the surface when *O* is not moving his hand; *E* should start the surface movement *after* *O* starts the hand movement and finish the surface move-

*Received in the Editorial Office on November 16, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

ment *before* *O* finishes. *O* should be instructed to move his hand in a continuous motion.

A variation of the above illusion is a length illusion in which moving a smooth surface in the same direction as the blindfolded observer's hand movement causes the surface to feel longer, and moving a surface in the direction opposite to the observer's hand movement causes the surface to feel shorter. In moving the surface in the same direction as the hand movement (the lengthening illusion) *E* must be careful not to exceed *O*'s rate of hand movement at any time, or else *O* will feel the surface moving in the wrong direction under his finger; i.e., he will feel the surface moving in the direction his finger should be moving. For the length illusion, as for the curvature illusion, *E* should not move the surface when *O* is not moving his hand. Sometimes, to obtain the length illusion it may be necessary to eliminate auditory cues arising from the friction between the surface and finger by which the observer may judge the rate of movement of his finger relative to the surface. A smooth surface should be used for the length illusion, because a change in relative rate of movement of hand and surface seems to be easier for an observer to detect with a rough surface than with a smooth one.

170½ West Frambes Street
Columbus, Ohio

REGRESSIVE BEHAVIOR CHANGES IN THE TUBERCULOUS PATIENT*

Department of Psychology and Psychiatry, Catholic University of America

SOL CHAREN¹

A. THE PROBLEM

In the field of phthisiotherapy, physicians have long maintained that the enforced regime of bedrest and hospitalization often creates behavior changes in patients (6, 13, 30, 35, 39) since normal social living and planning are disrupted and patients' reactions to tuberculosis infection are often emotionally harmful. In part these beliefs are due to clinical impressions and in part to the hypothesis that tuberculosis is such a frustrating social disease that only the strong and capable can withstand the concomitant long hospitalization and conditions of dependency. The opinion about this disease is well stated by Barker, Wright, Meyerson, and Golnick who have investigated the general field of chronic illness and consequently believe that the actual helplessness of the ill person inevitably creates personality changes. They state:

1. The sick person's interests will become progressively *narrow*; fewer stimuli will exist for him and he will respond, other than perceptually, to fewer of them. This follows from the reduced scope of his objective physical and social world and the great potency of a few internal stimuli. The sick person becomes "provincial," relative to his previous "cosmopolitanism." This behavior is both descriptively and dynamically regressive; ill persons are childish in behavior for the same reason that children are.

2. The ill person will become relatively *dominating, intolerant, and selfish*. To the degree that he lives in a small highly egocentric world, he will be unable to appreciate the needs of others; his behavior will be determined by his own needs to the exclusion of others. . . .

3. The ill person will become *dependent and insecure*. Paradoxically, along with domination, intolerance and selfishness go dependence, insecurity, and fearfulness. . . . In this respect, too, his behavior

*Received in the Editorial Office on November 21, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹This report was composed for the fulfillment of the publication requirement for the degree of Doctor of Philosophy in the Graduate School of the Catholic University of America, Washington, D. C. I wish to express my appreciation and thanks to Dr. William D. Commins, my major adviser, for his guidance.

is not unlike that of a child who loves his parents for being good to him, and hates them because he is so helpless without them and so easily coerced by them.

4. The ill person will become hypochondriacal. The things of paramount importance to him are intensely personal—the things that happen within his own body and the routines that minister to his physical welfare. . . . His thought, conversation and preoccupations deal with his own symptoms. This again is regressive behavior (3, pp. 322-323).

The problem under investigation can be stated as follows: Do tuberculous patients, as a result of enforced bedrest and dependency on others, show the pattern of behavior described by Barker *et al.* which they list as examples of regression? If patterns of regressive behavior are precipitated by the enforced conditions of bedrest, dependency, and frustration, should not these patterns disappear when the patients are no longer invalid and again able to assume a more normal way of living?

B. METHOD

1. Subjects

Twenty-five male veterans, patients in a tuberculosis hospital for six or more months and diagnosed as having pulmonary tuberculosis, were selected as the experimental group. To protect against undue stress only those were chosen who were non-toxic, afebrile, and in satisfactory clinical condition, with roentgenographic and clinical improvement satisfactory to the ward physician. Such patients are classified as Class III by the hospital. Prior to reaching this stage of improvement these patients had been restricted daily for a period of 24 hours, then 18½ hours in bed. As Class III patients they spent about 15 hours in sleep and rest.

Two control groups were obtained. One was of tuberculous patients who had been through the same hospitalization as the experimental group and were now waiting for discharge. These controls were classified as Class V by the hospital and had had serial x-rays stable for at least three months, sputum smears negative for four months, and sputum cultures negative for three consecutive months. They had full recreational privileges, engaged in one to four hours of rehabilitation activity daily, but were restricted to three hours of bedrest during the day. Both this group and the experimental group were selected by ward physicians with the stipulation that subjects be patients with pulmonary tuberculosis without complicating additional illnesses. Both groups were patients at the Veterans Administration Hospital, Baltimore, Maryland.

The second control group was selected from the Surgical Service of the

Fort Howard, Maryland, Veterans Administration Hospital. Ward physicians there selected patients receiving surgery whose total hospital stay was not expected to exceed two months. This group represented what might be considered a hospital population of normal individuals, since they were hospitalized under conditions that should not encourage invalidism, were not under too severe stress, and differed presumably from the non-hospitalized general population only by virtue of temporary absence from home and jobs. Physicians were asked to exclude those patients who gave evidence of being either behavior problems or neurotic.

Finally, the experimental group was used as its own control since it was followed until each individual reached Class V, was ready for discharge, and was then retested. With two groups of Class V patients a check was thus possible against chance selection of atypical patients.

All groups were closely matched for age, education, and vocational level, as Tables 1 and 2 show. The mean age was between 33 and 35 years, edu-

TABLE 1
AGE AND EDUCATIONAL LEVEL OF ALL PATIENT GROUPS

Data and group	N	Mean	Sigma	Range
Age (Years)				
III-eg*	25	33.2	10.6	20-56
V-c**	25	32.9	10.5	20-59
S-ct†	25	35.4	10.5	20-57
Education (Grades)				
III-eg		10.5	2.0	6 to 3 years of college
V-c		10.5	2.4	6 to 2 years of college
S-c		10.6	2.4	6 to 2 years of college

Note.—Differences between means are not statistically significant.

*Experimental group of Class III tuberculous patients.

**Control group of Class V tuberculous recovered patients.

†Control group of surgical patients.

TABLE 2
OCCUPATIONAL CLASSIFICATION OF PATIENT GROUPS

Description of occupation	Patient group		
	III-eg* No.	V-c** No.	S-ct† No.
0-X Professional, Technical and Managerial	3	4	2
1-X Clerical and Sales Work	5	5	4
2-X Service Work	4	4	4
4-X Mechanical Work	8	8	10
6-X Manual Work	5	4	5

*Experimental group of Class III tuberculous patients.

**Control group of Class V tuberculous recovered patients.

†Control group of surgical patients.

cation averaged middle tenth grade, with the range for age and education identical. Specific occupations were subsumed under the basic classifications of the Dictionary of Occupational Titles, Part IV (16) and, as Table 2 shows, little difference existed between any groups, although the surgical controls had a few more individuals in jobs which required manual effort. Each group had 17 white and eight Negro patients. All were veterans who had served in World Wars I and II and the Korean Campaign. No attempt was made to match for war service since the age matching in a sense covered this. The mean length of hospitalization for the experimental group of tuberculous patients was 8.3 months, for the tuberculous control group 12.4 months, and for the surgical controls 26 days with a range of 10 to 66 days.

2. Tests

Testing was done by group methods using both objective type paper-and-pencil tests, the Blacky, and the Rorschach. The tests and the purposes for which they were used are as follows:

a. Paper-and-pencil tests.

(1). *To measure the narrowing of interests.* The Strong Vocational Interest Blank for Men, revised (VIB) (36) was used in order to obtain a measure of the following interests: (a) Occupational interests, (b) Amusements, (c) Activities or hobbies, (d) Maturity of interests, (e) Level of occupational interest.

To obtain the last two interests (d and e) the Strong Interest-Maturity and Strong Occupational Level Scales were used directly. High scores on the former measure maturity of interest; high scores on the latter show how high up the occupational ladder one's interests go.

For the first three interests listed (a, b, c) the usual scale method of measurement of the Strong was not used; rather, Hofstaetter's "Actuality Measure" was adapted as a means of measuring the intensity of interest or ego involvement these subjects had in hobbies, amusements, work (23). Part I of the Strong lists 100 occupations which subjects check as "Liked," "Indifferent to," or "Disliked." These were selected as permitting a measure of the patients' interest in jobs. Part III of the Strong lists 49 items which deal with amusements and are answered as in Part I. This was used as a measure of interest in amusements. Last, Part IV under the heading "Activities" lists 48 items referring to hobbies and part-time interests.

Hofstaetter's "Actuality Measure" makes possible some quantitative determination of the amount of ego involvement present in any question or topic. The more involved a person or group is in a question, the greater

the tendency to take a "yes" or "no" stand and to avoid "no opinion" (23). The lower the "no opinion" vote the higher the "Actuality Measure." Lack of interest and consequent checking of the "no opinion" box (in this research, the "Indifferent to") leads to a low "Actuality Measure."

In addition to the five measures of interests another device was employed, determining whether patients were still in touch with current events. Part VIII of the Personal Audit Test of Adams and Lepley (Form *LL*) was selected as a current events test since it contains items dealing with social, political, and economic events of current interest (1). Again Hofstaetter's "Actuality Measure" was used since the degree of interest of patients in such questions was desired.

(2). *To measure the sick person's tendency to become dominant, intolerant, and selfish.* Six tests were used to measure this area.

(a). *An adaptation of the Dymond Empathy Scale* (17). Dymond defined empathy operationally as the difference between an individual's rating of a subject and the subject's self-rating. There are four parts: in Part 1, the subject rates himself; in Part 2, he rates another person; in Part 3, he predicts how the other will rate himself; in Part 4, he predicts how the other will rate the subject. The deviation score obtained is the number of points the subject is in error in matching his own Parts 3 and 4 against the partner's 1 and 2 ratings. Dymond found agreement between such scores and judges' ratings, Rorschach variables, and self-ratings. Low scores represent a high degree of empathy; high scores, the opposite.

Because her traits seemed too difficult of comprehension for a non-college group Dymond granted permission to modify her scale² and the one finally used was one adapted by Lorr from another developed by Brownfain (9).

(b). *A tolerance-intolerance scale.* This too was obtained from the Personal Audit Test (Form *LL*) of Adams and Lepley (1). Part VI of this test is a list of persons with very undesirable characteristics: the subject answers whether he dislikes such persons "Much," "Some," or "Not at all." A high score on this test gives some idea of the degree of irritability with, or inability to tolerate, disturbing traits in others.

(c). *The prejudice scale (Pr)* as a measure of intolerance, selfishness, and related traits. The test used is one adapted by Gough from Minnesota Multiphasic Personality Inventory (*MMPI*) items. Gough states that his items have value in determining prejudice or bias. He obtained high correlations with the E-F scales of the California Public Opinion Study;

²Personal communication, May 13, 1954.

further he found differences between two groups of students on a battery of tests which seemed to determine that the *Pr* test measured prejudice (18). For this test *T*-scores are obtained and high scores are expected of those most prejudiced.

(d). *The ascendancy-submission test.* This well known test by G. W. and F. H. Allport measures the disposition of the individual to dominate his fellows or to be dominated by them in face-to-face situations (2; 10, pp. 51-52).

(e). *Two devices were used to measure hostility.* First, patients scored themselves on this trait, and second they were rated by ward nurses. The test used by patients was the Moldowsky Hostility Scale consisting of items from the *MMPI* and validated by Dinwiddie who had psychiatrists in a clinic rate their veteran outpatients for the presence of hostility; he obtained high correlations between this criterion and clinic patients scores on the Moldowsky (15).

The second hostility measure was a rating scale of three items which Lorr, Rubenstein, and Jenkins obtained when they factor-analyzed a rating scale for measuring the personality of outpatients in mental hygiene clinics. One of the factors found was called Factor *H*, a general and first-order factor of hostility. The syndrome found was that of violence, temper tantrums, quarrelsomeness, etc. (27). In using Factor *II* as a rating scale, nurses were first given instructions and practice until correlations between nurse-judges was above .80. As only a total of seven nurses was used as judges, two of whom did over 60 per cent of the ratings, the training was not difficult. Unfortunately, patients at the Fort Howard, Maryland, Hospital were not rated because of hospital regulations.

(3). *To measure the degree of dependency and insecurity.* Two tests were used to measure this area:

(a). *The Maslow Security-Insecurity Index* (28). Maslow, Hirsh, Stein, and Honigman developed three 25-item tests designed to measure security-insecurity. The test was validated on student groups of 360 persons by self-ratings of students seeking guidance. Correlation between this test and the Thurstone Neurotic Inventory is .68.

(b). *The Winne Scale of neuroticism* (38). The Winne, like the Taylor Scale of Anxiety (37), is made up of items from the *MMPI*. Deese *et al.* found a correlation between the two tests of .668 with a population of 227 college students (14). Winne found 30 items of the *MMPI* which distinguished two groups at the .01 level of significance. One group was

of diagnosed neurotics and the other of normal subjects. Cross validation revealed the same degree of correlation.

(4). *To measure the degree of hypochondriasis.*

(a). *The Hypochondriasis Scale of the MMPI.* Because this scale has good reliability in measuring the degree of hypochondriacal concern and has established validity (20), it was the only test used to measure this area of regression. In this research this test was used as a separate, since Charen (12) found that almost identical scores were given by subjects whether the entire *MMPI* was given or only the hypochondriasis items.

b. *The Rorschach Test.* This well known inkblot test has measures which can be used to determine the presence or absence of regressive behavior changes. Thus Hertzman, Seitz, and Orlansky were able to study the stability of personality structure under anoxia by means of the Rorschach (22). The literature on this test, moreover, contains sufficient references to personality variables quite similar to those being employed in this research. The Rorschach patterns used to check and to complement the paper-and-pencil research will be described separately with the results obtained.

c. *The Blacky Test.* This test consists of 12 cartoons which portray the adventures of a dog named Blacky in situations which fit psychoanalytic theory. As the Blacky test has been fully described elsewhere (7) and is well known it will not be discussed here. Scoring of the test was objective using only the multiple choice questions which accompany each cartoon, and using Blum's criteria (8). Certain of the items in each group of questions, he states, are significant for the area under investigation, and if chosen by the subject are supposed to mean he has trends towards the Freudian area investigated by the particular cartoon. Using Blum's terminology, the subject may be "Strong" or have a definite weakness of psychosexual development in the specific area; or he may be "Weak" and have an absence of such regression.

3. Statistics

Analysis of variance was used in the comparison of groups and the statistical significance sought was the .05 level. For the Rorschach test normality could not be assumed. Instead, the median score for all groups combined was obtained and chi square used to determine the number in each group above the total median. The same procedure was applied to the results of the Blacky test as to the Rorschach, and the chi square statistic used to compare groups. Finally, for those tests where the "Actuality Measure"

was obtained, the chi square statistic was also used since normality could not be assumed.³

C. RESULTS OF THE PAPER-AND-PENCIL TESTS

On the 15 paper-and-pencil tests given to measure the possible presence of regressive structural personality changes in tuberculous patients, only two gave mean scores which brought out a significant difference between the experimental group and the three control groups. Since these two test means did not indicate the experimental group (III-eg) was regressed there is no evidence from the paper-and-pencil tests to support the hypotheses under investigation. Actually, the experimental group had more of its members interested in current events and showed less indications of dominance. For the remaining 13 tests mean scores did not differ significantly either for the experimental group compared to the two control groups, or for the experimental group compared against itself when retested after recovery from the disease.

To determine whether significant differences might exist between tuberculosis patients as compared to surgical patients, test scores of both the experimental (III-eg) and control group (V-c) tuberculosis patients were combined to form a single group (T) and scores were then compared to the surgical control group (S).

Neither the *F* ratio nor use of chi square revealed any significant differences. Long term hospitalization, over six months in the case of the one part of the *T* group, and 12 months in the case of the other, did not bring out deviant trends in personality structure as measured by tests and comparison with a short-term hospitalized group of surgical patients.

One interesting fact was, however, noted. Differences between group means was usually one of two or three points and not statistically significant, yet in 13 of the 15 tests the *T* group was consistently in the direction of unfavorable adjustment compared to the *S* group. This finding might be offered as evidence that long hospitalization and disease chronicity affect the tuberculous patient, perhaps in terms of emotional reactions rather than change in personality structure.⁴

³Personal communication, Dr. Peter Hofstaetter, July, 1954.

⁴To save printing costs a six-page table giving data for results discussed in the article has been deposited with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Order Document No. 4752, remitting \$1.25 for photoprints or \$1.25 for 35 mm. microfilm. Make checks or money orders payable to: Chief Photoduplication Service, Library of Congress.

D. RORSCHACH TEST

1. *Variables Selected*

One criticism of paper-and-pencil tests is that subjects can consciously or unconsciously try to present themselves in an unfavorable light (11). Projective or unstructured tests, because they do not require an opinion, are not answered in a yes or no manner, can avoid the problem of this self-glorification tendency although such variables as the test situation and the examiner can influence results (26). The use of the Rorschach in this research suggests an interesting comparison of the two types of devices.

The group Rorschach was used according to direction (19) but each card was presented for two minutes only and not turned. Beck's scoring and norms were used, but animal movement (FM) and movement caused by inanimate forces (m) were also tabulated. To avoid the shading controversy of the various Rorschach schools (21), all chiaroscuro responses were given one scoring depending on the presence and importance of form. Shading scores were then F(c), (c)F, and (c) following Kendig (24). All records were scored by this writer. To check his accuracy 10 records were taken at random from the total number of Rorschach protocols and rescored by an experienced clinical psychologist and a fourth-year graduate student in clinical psychology.⁵ Only a five per cent discrepancy was found by these two judges in the 10 records which contained a total of 183 responses.

a. *To measure the narrowing of interests.*

(1). *Populars.* Rorschach emphasized that the popular or "vulgar" responses measure the individual's ability to share in the collective or common way of sensing or perceiving (34, p. 198). Beck's statistics show that Populars (P) lower than four in an adult show "we are dealing with one whose thinking fails in identification with that of the group" (4, p. 17).

(2). *Animal per cent.* The animal or A response Rorschach saw as the indicator of stereotyped thinking and inversely related to freedom of association and fluidity of thought processes (34, p. 68). High A per cent means less ability to have a variety of ideas and varied interests. Beck sets the upper optimal range of A per cent at 65 (4, p. 15).

(3). *Human content.* References to humans in the Rorschach test Beck sees as an indication of interest in all known fields, "the dance, the dramatic arts, sports and adventure, war, religion and other rituals, mythology, . . . abstractions of human qualities, humans in meaningful scenes, carica-

⁵Dr. Shabsie Kurland and Mr. Louis Lepine kindly consented to act as judges.

ture" (4, p. 43). The cutting-off point was set at one or no human (H) response (this was done because the human of Card III is so readily seen).

(4). *Number of responses.* A fourth indicator of interests is the number of responses (R) which Beck interprets as one measure of the freedom to think and an index of the individual's ability to have a variety of ideas and associations (4, p. 54). Beck furnishes norms for R but his figures are based on various *IQ* groups and could not be used in this research. Instead the median R for all four groups was used.

(5). *Variety of content.* Finally, the variety of content in each record is an indicator of interests (5, p. 28). Human, animal, and anatomy references were not considered in determining content variety.

b. To measure dominance, intolerance and selfishness.

(1). *Color.* Rorschach described the form element in his test as measuring conscious control (34, p. 69). Piotrowsky states that the more dominant this is over color, the greater the control over emotional impulsivity. Form color (FC) thus stands for effortless and adequate emotional adaptation and color (C) for emotional impulses in which no regard is given to the rights and wishes of other persons and in which self-centeredness and lability are highest; color form (CF) responses are indicators of reactions between these two (31). It is thus desirable that an adult produce more FC than CF and C (32, p. 343). The reverse is indicative of immaturity of emotional reactions.

(2). *Sum color.* Both Beck (4, p. 31) and Piotrowski (32, p. 343) regard the sum of all color as a measure of the individual's desires to associate or dissociate with people, of the capability of feeling content with his world. The lower the sum color (Sum C) the less is this desire for contact (31).

(3). *Movement.* Movement or seeing humans in some kind of activity on the Rorschach plates represents among other things a value system which is mature (M); seeing animals in motion (FM) and seeing motion explained as due to natural causes (m) are both concepts of the self which the individual has and which are less mature (25, pp. 254-291). Piotrowski believes that in moments where the individual loses control over his behavior the effect of FM increases (32, p. 356). The mature person has more M than the total of FM and m combined. Illness, if it creates regressive changes, should result in the Rorschach picture of more FM and m combined than of M.

(4). *Experience-balance*. Beck states that the ego defends itself against frustration by a process of withdrawal or coercion and this is directly measured by the Rorschach sign of the ratio of color responses weighted to movement responses (M), or the experience-balance (EB) (5, pp. 32, 40). Rorschach has defined coercion as resulting in an EB of one M:one C, or one M:no C, or no M:no C (34, p. 84).

(5). *FM of hostile animals*. Finally, if persons are egocentric we can expect them to be easily irritated and prone to have hostile attitudes. The sign of animal movement (FM) is the measure of less mature concepts and one would then expect the irritated individuals to see animals that are hostile or in threatening positions (32, pp. 356-357).

c. *To measure the degree of dependency and insecurity.*

(1). *Shock*. Both subjective and objective signs of disturbance to the chromatic or achromatic Rorschach plates are considered indications of emotional instability or neuroticism. The criteria for these signs have been thoroughly established in the literature (25, pp. 338-349).

(2). *H:Id*. The ratio of the number of parts of the human body seen (Hd) to the total persons seen (H) has long been accepted empirically as a measure of anxiety when the Hd responses outnumber the H responses by a two-to-one ratio.

(3). *Vague wholes*. As a further measure of insecurity and dependency, the number of vague whole responses seen is supposedly a good yardstick. A vague whole response is one which uses the entire inkblot, is of such poor quality that it does not meet acceptable standards (F—) or which is of a percept which is vague and similar to the concept of "inkblot," such as "cloud" or "smoke" or other amorphous perceptions. Whole responses measure the abstracting, surveying, and integrative abilities of the individual. Where this is impaired or weak we may assume inability to analyze and abstract successfully, and anxiety can be deemed present (33, pp. 138-140).

(4). *W:M*. Another index of dependency may be used, namely the ratio of all whole responses to movement responses (W:M). This ratio normally should be two to one and is believed to be the ability or extent to which the individual is able to mobilize his creative or productive energies to back up his intellectual interests and ambitions (25, p. 302). When the ratio is higher than two to one, the level of aspiration is too high and the drive can become a negative one. The individual strives hard but accomplishes little, or if he does achieve, this is at the expense of worthwhile development of personality.

(5). *Shading.* Shading is an index of insecurity or anxiety. Much of what has been said about the difference between the different kinds of color responses can be applied in this discussion of the shading responses. Where the form element is dominant over the shading (Fc) there is greater control over anxiety than when the form element is either lacking (c) or between the two (cF) (31). Thus where the shading with form prominent is less than the combined score of the other two types of shading we can assume that anxiety is not well controlled.

d. *To measure the degree of hypochondriasis.* The measure of hypochondriasis or body preoccupation on the Rorschach is the anatomy response.

2. Results of Rorschach Test

Little difference between the experimental group of tuberculous patients and its controls were found on the results of the Rorschach analyses. The experimental group did show changes in the test-retest situation which were not consistent but none of these changes were statistically significant nor was there any Rorschach evidence which showed the experimental group basically different from the other two control groups.

In a further effort to emphasize differences the experimental group results on first testing were combined with the tuberculous control group (V-c) to form one tuberculous group (T) and these results compared to the surgical control group test results. Some differences between the T group and the surgical patients (S) were then obvious. Notice, under the heading, "Narrowing of Interests" in Table 3, while only one of five Rorschach measures shows a statistically significant difference between the groups, there are definite differences in the means of the remaining four measures which show the T group as having Rorschach evidence for such narrowing.

The same kind of trend is also obvious under the second heading in Table 3, "Dominating, Intolerant, etc." While only one of the five Rorschach measures is statistically significant and then only at the .10 level of significance, there are obviously more in the T group on four of the five measures who show Rorschach evidence for regressive emotional changes.

The third and fourth headings of Table 3 show a reverse tendency in that there the surgical control group has more members with Rorschach signs of dependency, insecurity, and hypochondriasis. Combining these two headings, we find that in three of the six Rorschach measures statistically significant differences are found between groups which show the surgical control group more maladjusted.

A possible interpretation is that the surgical control group may have among

TABLE 3
NUMBER AND PER CENT OF TUBERCULOUS AND SURGICAL PATIENTS WITH RORSCHACH
SIGNS OF REGRESSION

Rorschach sign	N =	Combined tuberculous 50		Surgical 25		Level of confidence
		No.	%	No.	%	
<i>Narrowing of Interests</i>						
P less than 4		7	14	4	16	not significant
A% over 65		19	38	5	20	not significant
H of 1 or less		27	54	11	44	not significant
R*		28	56	12	48	not significant
Content*		30	60	7	28	.01
<i>Dominating, Intolerant, etc.</i>						
CF, C over FC		20	40	5	20	.10
Sum C*		24	48	9	36	not significant
FM, m over M		17	34	12	48	not significant
Constricted EB		18	36	8	32	not significant
FM of hostile A		9	18	4	16	not significant
<i>Dependency, Insecurity</i>						
Neurotic shock		24	48	19	76	.02
Hd over H		5	10	4	16	not significant
Vague W		28	56	17	68	not significant
W over M		29	58	20	80	.06
Shading		20	40	11	44	not significant
<i>Hypochondriasis</i>						
Anatomy of 1 or more		17	34	15	60	.05

*Criterion is the number below the median score for both groups.

its members more insecure, neurotic individuals (Neurotic Shock) who are greatly concerned about surgery (Anatomy Responses), do not feel capable of the demands made of them (W:M ratio), but in terms of attitudes to others and of interest in their normal social environment they still think of themselves as part of the extramural environment. It is interesting that in the one Rorschach sign which Piotrowski believes appears during moments where the individual loses control over his behavior, namely FM (32, p. 356), we find that in Table 3 the surgical group has 48 per cent who have an excess of FM and m over M. In the combined tuberculous group only 34 per cent show this ratio. There is thus in the surgical group Rorschach evidence of insecurity and fearfulness, concern about the body. This evidence is not as marked in the tuberculous group of patients although they give evidence of being more "intolerant and selfish" and of becoming responsive to "fewer stimuli" (3, pp. 322-323).

In contrast with the surgical patients the tuberculosis patients have fewer who show Rorschach evidence for body concern (Anatomy) or who are in-

secure and neurotic. As of the period when tested the tuberculosis patients seem to have accepted their illness well and to have adjusted to it, although they manifest some mild changes in the emotional sphere. As with the paper-and-pencil tests the Rorschach results do not yield statistically significant differences. Only the term "tendency" can be applied, but within the limits of this word there is Rorschach evidence that both tuberculosis and surgical groups do show some emotional reactions to hospitalization. The former give evidence of good adjustment to their situation, but some unfavorable behaviorisms toward their social environment seem possible. The surgical group gives test evidence for being outwardly adjusted and inwardly under subjective strain.

E. THE BLACKY TEST

Unfortunately for this part of the study technical difficulties made it impossible for this test to be given to the surgical control patients at the Fort Howard Veterans Administration Hospital. This section therefore deals with the differences between the experimental group and the tuberculous control group, the former being tested and then retested after recovery.

The results of this test do not support a hypothesis of regression; in fact, the experimental group, while ill, shows more evidence of "Weak" tendencies than when recovered. When the experimental group did recover it did not resemble the matched group of tuberculous patients, rather it showed statistically significant differences. These differences suggest that if the Blacky test is valid some unknown variable is influencing the test picture.⁶

F. SUMMARY

The hypothesis was made by Barker, Wright, Meyerson, and Golnick (3) that the chronically ill patient tends to regress in four characteristic ways. By means of paper-and-pencil tests, the Rorschach, and the Blacky test an attempt was made to measure this. Hospitalized tuberculous patients were the research subjects since the chronicity of this illness requires long periods of bedrest, dependency on hospital personnel, and in general creates stress.

Twenty-five male veteran patients sick with pulmonary tuberculosis with an average hospital stay of eight months were matched for age, education,

⁶Further investigation revealed low reliability of the Blacky test caused these inconsistent changes. See Charen, S. Reliability of the Blacky Test. *J. Consult. Psychol.*, in press.

color, and occupational level first with 25 recovered tuberculosis patients with an average hospital stay of 12 months; and second with a group of 25 non-tuberculosis patients hospitalized less than two months for surgery.

1. *Paper-and-Pencil Test Results*

No significant differences were found between any of the groups on these tests measuring four different kinds of attitudes believed to exist in the chronically ill, nor did the experimental group show any test change when recovered and retested. Combining the two groups of tuberculosis patients and comparing tests for this new group with the test results of the surgical group revealed insignificant but consistent test evidence for the presence of minor effects of hospitalization and illness in the tuberculosis patients.

2. *Rorschach Test Results*

No significant differences were found between any of the groups on Rorschach patterns measuring the four kinds of attitudes. Combining the two groups of tuberculosis patients and comparing the Rorschach patterns for this new group against the surgical control group revealed tendencies, as in the paper-and-pencil tests. The surgical patients seemed more concerned about surgical intervention since the Rorschach patterns revealed more signs of insecurity, neurotic shock, and hypochondriasis, but they showed Rorschach evidence for better outward social adjustment. The tuberculosis patients were emotionally more secure, but they showed tendencies towards narrowing of interests, intolerance, and selfishness.

3. *The Blacky Test Results*

In this measure for Freudian concepts no evidence for regression in the experimental group could be found. Further investigation showed poor reliability in this test caused conflicting results.

In none of the tests was there evidence for changes in personality structure in any of the research groups to which the term regression could be applied. Tuberculosis patients on the average apparently accept hospital conditions of bedrest, dependency, reliance upon others, and social frustration with patterns of behavior which are adult and not childlike.

REFERENCES

1. ADAMS, C. R., & LEPLEY, W. M. *The Personal Audit (Test)*. Chicago: Sci. Res. Assoc. (no date).
2. ALLPORT, G. W., & ALLPORT, F. H. *The A-S Reaction Study: Manual of Directions*. Boston: Houghton Mifflin, 1939.

3. BARKER, R. C., WRIGHT, B. A., MEYERSON, L., & GOLNICK, M. D. Adjustment to Physical Handicap and Illness: A survey of the Social Psychology of Physique and Disability. New York: Soc. Sci. Res. Council, Bull 55, Revised, 1953.
4. BECK, S. J. Rorschach's Test: Vol. II. New York: Grune & Stratton, 1949.
5. ———. Rorschach's Test: Vol. III. New York: Grune & Stratton, 1952.
6. BROWN, L. Mental aspects in etiology and treatment of pulmonary tuberculosis. *Internat. Clinics*, 1953, **3**, 149-174.
7. BLUM, G. S. A study of the psychoanalytic theory of psychosexual development. *Genet. Psychol. Monog.*, 1949, **39**, 3-102.
8. ———. Revised Scoring System for Research Use of the Blacky Pictures: Male Form. Mimeographed by the author, 1951.
9. BROWNFAIN, J. J. Stability of the self-concept as a dimension of personality. *J. Abn. & Soc. Psychol.*, 1952, **47**, 597-606.
10. BUROS, O. K., Ed. The Fourth Mental Measurements Yearbook. New Jersey: Gryphen Press, 1953.
11. CHAREN, S. The Rorschach test. *Bull. Med. Library Assoc.*, 1953, **41**, 208-214.
12. ———. A note on the use of a paper-and-pencil form of the MMPI scale for hospital use. *J. Consult. Psychol.*, 1954, **18**, 344.
13. COLEMAN, J. V., HURST, A., & WHEDDER, J. L. Tuberculosis and dependency. *Psychosomat. Med.*, 1950, **12**, 189-198.
14. DEESE, J., LAZARUS, R. S., & KEENAN, J. Anxiety, anxiety reduction, and stress in learning. *J. Exper. Psychol.*, 1953, **46**, 55-60.
15. DINWIDDIE, F. W. An application of the principle of response generalization to the prediction of aggressive responses. Unpublished doctors dissertation, Catholic Univ. of Amer., 1954.
16. DIVISION OF OCCUPATIONAL ANALYSIS, WAR MANPOWER COMMISSION. Dictionary of Occupational Titles, Part IV. Washington: U. S. Govern. Print. Office, 1944.
17. DYMOND, R. F. Personality and empathy. *J. Consult. Psychol.*, 1950, **14**, 343-350.
18. GOUGH, H. G. Studies of social intolerance: II. A personality scale for anti-semitism. *J. Soc. Psychol.*, 1951, **33**, 247-255.
19. HARROWER, M. R., & STEINER, M. E. Large-Scale Rorschach Technique. Springfield, Ill.: Thomas, 1945.
20. HATHAWAY, S. R., & MCKINLEY, J. C. Manual for the Minnesota Multiphasic Personality Inventory. New York: Psychological Corp., 1951.
21. HERTZ, M. The shading response in the Rorschach inkblot test: A review of its scoring and interpretation. *J. Gen. Psychol.*, 1940, **23**, 123-167.
22. HERTZMAN, M., SEITZ, C. P., & ORLANSKY, J. Stability of personality structure under anoxia. *J. Gen. Psychol.*, 1955, **52**, 65-73.
23. HOFSTAETTER, P. The actuality measure in the study of public opinion. *J. Appl. Psychol.*, 1953, **37**, 281-287.
24. KENDIG, I. V. The basic Rorschach triad. *J. Proj. Techn.*, 1954, **18**, 448-452.
25. KLOPFER, B., AINSWORTH, M. D., KLOPFER, W. G., & HOLT, R. R. Development in the Rorschach Technique: Vol. I. Technique and Theory. New York: World Book, 1954.
26. LORD, E. Experimentally induced variations in Rorschach performance. *Psychol. Monog.*, 1950, **64**, No. 316.
27. LORR, M., RUBENSTEIN, E., & JENKINS, R. L. A factor analysis of personality ratings of outpatients in therapy. *J. Abn. & Soc. Psychol.*, 1953, **48**, 511-514.
28. MASLOW, A. H., HIRSCH, E., STEIN, M., & HONIGMANN, I. A clinically derived test for measuring psychological security-insecurity. *J. Gen. Psychol.*, 1945, **33**, 21-41.

29. MCNEMAR, Q. *Psychological Statistics*. New York: Wiley, 1949.
30. PETERSON, B. Mental attitudes as a factor in the treatment of tuberculosis. *Amer. Rev. Tuberculosis*, 1929, **17**, 383-393.
31. PIOTROWSKI, Z. A. *A Rorschach Compendium*. Utica, N. Y.: State Hospital Press, 1950.
32. PIOTROWSKI, Z. A., & SCHREIBER, M. Rorschach perceptanalytic measurement of personality changes during and after intensive psychoanalytically oriented psychotherapy. In: G. Bychowski & J. L. Despert (Eds.), *Specialized Techniques in Psychotherapy*. New York: Basic Books, 1952.
33. RAPAPORT, D. *Diagnostic Psychological Testing: Vol. II*. Chicago: Yearbook, 1946.
34. RORSCHACH, H. *Psychodiagnostic* (4th ed.). Berne: Huber, 1949.
35. STRECKER, E., & BRACELAND, G. B. Mental attitudes of tuberculosis patients. *Ment. Hyg.*, 1938, **22**, 529-543.
36. STRONG, E. K., JR. *Vocational Interests of Men and Women*. Stanford: Univ. Press, 1948.
37. TAYLOR, J. A. A personality scale of manifest anxiety. *J. Abnorm. & Soc. Psychol.*, 1953, **48**, 285-290.
38. WINNE, J. F. A scale of neuroticism: An adaptation of the Minnesota Multiphasic Personality Inventory. *J. Clin. Psychol.*, 1951, **7**, 117-122.

200 Rhode Island Avenue, N. E.
Washington 2, D. C.

A COMPARISON BETWEEN A HOSPITALIZED AND A NORMAL GROUP ON THE X-O TESTS*

Counseling Psychology Service, V. A. Hospital, Chillicothe, Ohio

A. W. JONES

A. THE PROBLEM

The change in values over the last 30-year period and a comparison of the attitudes of older and younger persons have been the subject of a recent investigation (2). At that time it was demonstrated that since 1923 there have been certain liberalizations in attitude, as revealed through a decline in the total number of words considered wrong on the X-O Form, and an increase in the number of words marked as interests. Little change in the total number of worries over the three-decade period was interpreted as indicative of less involvement of worries in the total cultural pattern. The older adults were found to have retained, in the main, the attitudes of their youth, being more resistive to cultural change, rather than growing increasingly more conservative with age. Percentage item analyses pointed up the increased freedom in sex-social activities among younger groups, together with a general relaxation of moral codes; and some liberalizing was observed among the older persons as well.

Surveying of the literature reveals that neuropsychiatric patients have been the subjects of only certain limited types of attitudinal research (1). With this background of data and accumulated groups for comparative purposes, it seemed unusually appropriate to see in which ways mental patients might differ in response on such an inventory as was previously utilized with normals at decade intervals over an extended period of cultural change.

B. MATERIALS AND CASES

Parts 1 and 2 of the X-O inventory were used in this study (3). The first test consists of a list of 125 possible wrongs, such as "fear, anger, suspicion, laziness, contempt," with the person taking the test directed to cross out any of these items he believes are wrong and then to circle the one word in each of the 25 lines thought worst. The second test consists of possible worries as "sin, headache, fault-finding, sneer, depression." All words ever worried about are to be crossed out and the one word most

*Received in the Editorial Office on November 25, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

worrisome in each line to be circled.¹ The third *X-O* test is concerned with a list of possible interests, but since it was not given to the adults in the previous study, it was also not administered to the hospitalized group.

One hundred male chronic, open-ward patients at the Chillicothe, Ohio, V. A. Neuropsychiatric Hospital were the subjects. They were gleaned from a group of over 300 contacted patients and thus represent those who were able to read, were coöperative, and in sufficient contact to complete the form. The range in age was from 20 to 70, and the average 47. There were thus nine patients 20-29; 27 from 30-39; 21 from 40-49; 23 from 50-59; and 20 who were 60 and over. By far the largest number were diagnosed schizophrenic (70 out of 100), and there were only five who were classified other than psychotic. The educational level of the group was believed to be equivalent to the completion of grammar school, an average below that of the mostly college trained normals to whom they were compared.

These patients could be contrasted with a total of 278 males obtained in the 1953 testing with normals (2). (The larger groups of undergraduate and adult females are *not* included in this total). These consisted of 132 male undergraduate students at Ohio State University and 146 adult men. Of the latter, 32 were in their thirties, 36 in their forties, 40 in their fifties, and 38 were 60 years of age and over.

C. RESULTS

The two major methods of treating the data are: (a) analysis of the total number of words marked on each test; and (b) item analysis of the percentage marking each word on both tests.

1. *Age and Normal-Hospitalized Comparisons in Total Words Marked*

Table 1 presents the results as regards total number of words marked as wrong and as worries by both the patients and the normal adult males of the latest testing in 1953. Although in each instance there is an increase with age in average number of words marked on both tests, the normals mark significantly more (beyond .01 level) words only on the wrongs test. The previous study demonstrated the meaning of more frequent marking by each successive decade of adults on wrongs and showed that the difference between age groupings on the worries test, although increasing for the older cases, was not interpretable in terms of any set age trends.

¹The circling aspect is not dealt with herein.

TABLE 1
A COMPARISON OF MALE PATIENTS AND NORMAL MEN TESTED IN 1953 ON TOTAL WORDS MARKED ON THE WRONGS AND WORRIES SUBTESTS OF THE X-O TESTS

	Patients (N = 100)		1953 Normals (N = 278)	
	20-39	40 plus	20-39	40 plus
Age				
Wrongs	54.9	59.5	40.8	54.5
Worries	33.2	39.5	31.0	34.0

As compared to normals, patients show an insignificant increase in marking with age on the wrongs subtest. This quite small rise might possibly be due to the greater homogeneity of the patient sample (they all have the common experience of being in the hospital), which could reduce the appearance of any such age differences as might be present.

The slightly greater increase with age in marking on the worries test is also not large enough to be considered important. Again, as was discovered previously with the normals, worries appear to be more individualized in nature and less sensitive to gross age changes.

The second major finding is that with both tests and with both age groupings, the patients mark more words than the normals. On the wrong subtest, there are significantly (beyond .01 level) more words marked by the patients, indicating a more conservative, rigid attitude toward life in general. However, on the average significantly more worries are not marked by patients. Either they are unaware of their deviations or are unwilling to admit them in testing. A neurotic sample might more readily reveal symptoms than the primarily psychotic group reviewed here.

The patients of today mark about the same number of wrongs (average of 57) as the junior-senior students of 1923 (55), indicating a general framework of conservation appropriate to a period of 30 years ago, but quite different in content, as the item analyses will reveal. Since the worries test has undergone little change over the 30 years in total number of words marked, it is not surprising that today's patients, when viewed in toto, mark about 36 words, the same as the junior-seniors of three decades previous.

The diagnosis for each patient was obtained and breakdown of total score on the wrongs and worries tests was reviewed. The vast majority of cases carried the diagnosis of schizophrenia, and paranoid schizophrenia was, in turn, the most frequent type. There were no significant differences between paranoid schizophrenics and all other schizophrenics on total score. No attempt was made to compare psychotics to non-psychotics (neurotics, organics, etc.), since there were too few cases in the latter category.

2. *Item Analysis*

Because of the general absence of age trends within the hospitalized population, it was decided to "lump" all the patients together in the determination of percentage marking for each word. It was felt that by thus handling the cases more stabilized percentages would be obtained. Since the average age of the patient group was 47, it was compared with the normal males aged 40 and above. In this way, consideration would be taken of any age differences as might be present but did not become recognizable in the total score analyses.

On the worries subtest, there were a total of 40 words which showed at least 20 percentage points difference in marking between the normal and hospitalized groups. This amount of difference is equivalent to at least the .01 level of confidence. To facilitate the handling of these diverse differences, the 40 words were divided into predetermined categories by four judges, all of whom had either Ph.D. degrees or were advanced graduate students in psychology. The average percentage agreement of the word classifications was 85 per cent.

Table 2 shows the categories used and the words placed in each one. Only

TABLE 2
CATEGORIES OF ITEMS OF THE WORRIES TEST WITH AT LEAST 20 PER CENT POINTS
DIFFERENCE BETWEEN NORMAL AND HOSPITALIZED MALES: EXTENT OF AGREEMENT
OF JUDGES IS INDICATED

Neurotic complaints	Depression	Delusional thinking
100% Nightmares Hysterics Twitching 75% Fainting Suffocating Dizziness	100% Blues 75% Moodiness Ruin Crying	100% Persecution Enemies Poison 75% Burglars Lies
Physical impairment	Hostility Reaction	Immaturity
100% Disfigurement 75% Homeliness	100% Sneer Meanness 75% Roughness Faultfinding	100% Childishness Flightiness Giggling
Feelings of being different	Disease	Employment
100% Queerness Stupidity 75% Awkwardness	100% T.B. Germs 75% Dirt	100% Work Money Business Employer
Miscellaneous 100% Stylishness		

those words with at least 75 per cent judge agreement were included. For 35 out of the 40 words this level of agreement was obtained. *All* of the worries groupings were marked more frequently by the patients with the exception of *employment*, checked more often by the normals.

It can be seen readily that the greatest number of these differences fall into the general area of abnormal behavior. Certainly it might be expected that paranoid schizophrenics would be concerned about items classified as *delusional thinking*. *Neurotic complaints*, *depression*, *immaturity*, and *hostility reaction* can be similarly explained in that these symptoms are also characteristic of mental patients. The fact that *physical impairment* and *disease* are marked as worries is not surprising in a hospital setting. Quite interestingly, as indicated by the checking of words under the category, *feelings of being different*, patients would seem to be cognizant of the fact that they are unlike other people. It would not be anticipated that patients would be as concerned about *employment*, since the vast majority of them are not involved in work. The *miscellaneous* category, designed to be a "catch-all," had only the one word "stylishness," placed there with agreement, thereby reflecting the efficacy of the other predetermined categories.

To demonstrate the relationships between current percentage markings by both normal adults and patients, and current and previous markings by students, Table 3 was constructed. Here are listed in descending order those worries with the largest differences between patients and normals, with the

TABLE 3
LISTING OF WORRIES SHOWING THE GREATEST DIFFERENCES BETWEEN HOSPITALIZED AND
NORMAL CASES, PLUS THE ADDED PERCENTAGE COMPARISONS WITH
1923 AND 1953 JUNIORS-SENIORS

Word	'23 J-S %	'53 J-S %	Normal Adults %	N. P. %	Difference %
Poison	13	9	10	60	50
Giggling	7	7	8	54	46
Stupidity	21	24	27	69	42
Childishness	13	24	6	45	39
Insanity	15	16	16	54	38
Roughness	19	16	14	48	34
Suffocating	12	16	17	51	34
Twitching	6	10	12	45	33
Ruin	20	16	22	54	32
Teasing	14	20	11	42	31
Hysterics	6	5	5	36	31
Work	68	49	71	33	38
Employee	24	20	35	9	26
Business	36	16	52	30	22
Money	64	60	65	45	20

one addition at the bottom of the table of those words which were checked more frequently by the hospitalized group. The comparative percentages for each word are also listed for the 1953 junior-senior men and the 1923 junior-senior men.

The first two columns of the table show that the differences in percentages on these worries over the 30-year period are amazingly small, when viewed against the background of value changes previously reported (2). Of all the words listed, only "business" has changed 20 per cent points over the three decades. In addition, the percentage marked by the normal adults are also very close to these other two testings, as a glance at the third column will reveal. Again only the word "business" displays a 20 per cent point difference when the normal adults are compared to both the 1923 and 1953 juniors-seniors.² Finally, a totaling of all the percentage points differences between the patients and the 1953 adults on these 11 words amounted to 410 points; whereas the total difference in percentage points between the 1953 adults, the 1953 juniors-seniors, and the 1923 juniors-seniors combined was only 83. It therefore seems that normal and abnormal groups can be differentiated quite clearly on the basis of how the worries subtest is marked.

Six of the 10 categories are represented in these the greatest worries. The largest difference on the word "poison" does reflect the paranoid element in the hospitalized group. Those words included under the category "employment" are listed at the bottom of the table for comparative purposes and are self-explanatory.

The wrongs were found to be more difficult to categorize. At the 20 per cent point, there were 66 words which differentiated the hospitalized from the normal group. Therefore, 30 percentage points was taken as the required difference in order to obtain the most essential characteristics of the patient group. There were 39 words so designated, 38 of which were marked more often by the patients. Only "swiping" was marked more often by the normals as a wrong—91 per cent as opposed to the patients' 48 per cent. (It makes sense that hospitalized "vets" living side by side in a ward would not as often look askance at minor pilfering). As before, these words were placed into predetermined categories by the judges, but with somewhat less success. (The average percentage agreement was only 65 per cent). It appears that wrongs are less personalized in nature, more affected by cultural change, and thus may have less "power" as abnormal-normal discriminators.

²Quite naturally, older working males are more worried about business than younger students.

Only the various categories are here mentioned and various examples of the words herein contained examined: *hostility reaction*, *physical impairment*, and *feelings of being different*, are repeaters from the worries categories. Quite logically, the *self-assertive* category appears to be a mild form of *hostility reaction*, containing such words as "bluff," "boldness." The main wrong stressed in the *financial* group is "debt." The patients' economic dependence upon the government might be responsible for this reaction. *Depression*—"feeling worried," "blue," is quite an expected type of marking. In the *sex-social* area there is a crossing out of such terms as "flirting," "bashfulness," indicating a more rigid approach to heterosexual relations. "Dullness" and "absentmindedness" are such words as are included under the classification of *cognitive difficulties* and point to a condemnation of the "mental confusion" common among neuropsychiatric patients. A proper level of activity seems important, too, since both "overwork" and "idleness" appear as wrong under *work level*. *Visceral reaction* includes such terms as "overeating," and "vomiting," suggesting an awareness of bodily involvement in emotional disorders. *Socio-economic* status contains "peddler" and "aristocrat," somewhat rejecting economic extremes. Finally, the *miscellaneous* category presented no clear cut item groupings.

TABLE 4
LISTING OF WRONGS SHOWING THE GREATEST PERCENTAGE DIFFERENCES BETWEEN HOSPITALIZED AND NORMAL CASES, PLUS THE ADDED PERCENTAGE COMPARISONS WITH 1923 AND 1953 JUNIORS-SENIORS

Word	'23 J-S %	'53 J-S %	Normal Adults %	N. P. %	Difference %
Toughness	41	14	13	66	53
Quarrel	39	20	31	84	53
Insanity	36	11	19	69	50
Clumsiness	26	14	13	63	50
Flirting	47	11	18	66	48
Fretting	53	41	43	88	45
Slang	43	11	20	63	43
Vomiting	18	8	11	54	43
Giggling	44	17	18	60	42
Teasing	34	29	27	69	42
Bluff	43	26	41	82	41
Swiping	80	86	91	48	43

In Table 4 are found those wrongs which showed the largest percentage differences between the normal adults and the hospitalized patients, with the 1923 and 1953 undergraduate percentages listed for purposes of comparison. The total percentage points difference between the patients and

normals was 510, while that for the adults, the 1953 and 1923 undergraduates was 281, again bearing out the same kind of relationship found with the worries. As the table shows, however, the wrongs have changed considerably over the 30-year period; and adults and youths do differ more than was discovered with the worries. It has been indicated previously that much of this is because wrongs appear to be more sensitive to age and cultural change.

D. DISCUSSION

The uniqueness of this study seems to lie in the fact that it was possible to compare the "attitudes" of a neuropsychiatric group to those of normals of both today and 30 years ago. Such an investigation has great value in demonstrating that although patients in a mental hospital feel very differently and hold other standards from those which society calls normal, these feelings are in many ways akin to what was considered appropriate three decades previous! Indeed, psychotics may be caught, in part, in the "conservatism" of the past.

However, neither cultural nor age change appears to have had an appreciable effect on the worries of psychotics, for the words that differentiate them from normals have been marked by this latter group with amazing constancy for 30 years! Is it possible to conceive of a psychotic value structure which would essentially be resistive to changes? But other work must establish this.

As was mentioned, the hospitalized group did not display the age changes which were so typical of the normals. Certainly future studies along these lines should make an effort to obtain a broader sample than the chronic openward group tested here, and to examine differences in educational level. Patients do consider more things wrong, though, and reflect a "conservatism" which makes present day living strained and stultified. That more worries were not marked might be merely a general reluctance to admit to them; this needs further study.

The logicity of the item analyses makes extensive discussion of them unnecessary. It might prove useful to broaden the scope and increase the severity of the type of items included. The original *X-O* tests contained many such words, but they were later omitted when the test was used in schools. To be sure, there are certain objections to using the *X-O* as it stands, and these have been detailed elsewhere (4), but the use of such an item-check format with adult patient groups appears more than justified.

E. SUMMARY

One hundred chronic open-ward male patients at the Chillicothe V. A. Hospital were administered Tests 1 and 2 from the X-O form. Their responses were compared with those of 384 males obtained in the 1953 testing with normals, and, also with those of college upperclassmen of 1923. The following conclusions are offered:

1. The neuropsychiatric patients did not exhibit the increase in number of words marked with age on the wrongs test that had been shown previously with normals. Greater homogeneity of the patient sample was offered as a possible explanation.
2. Little difference in total number of worries marked by patients and normals was noted; and it was suggested that patients might not wish to admit to them.
3. Patients marked more wrongs than older normals to the level of "conservatism" used by college undergraduates of 30 years ago.
4. Item analyses pointed up certain general categories of response on both tests which differentiated patients from normals and which "made sense" in terms of the hospital setting.
5. The worries which separated patients and normals had shown little change in marking in 30 years, and in older adult and younger groups. It thus seems that patient groups could be clearly identified on the worries test.
6. The use of item-check type tests with psychiatric patients seems justified.

REFERENCES

1. FRENCH, E. G., & YARNOLD, J. K. The relationship between attitude survey scores and neuropsychiatric breakdown. *USAF Hum. Resour. Res. Cent., Res. Bull.*, No. 52-38, V. 9 p.
2. PRESSEY, S. L., & JONES, A. W. 1923-53 and 20-60 age change in moral codes, anxieties, and interests as shown by the X-O Tests. *J. of Psychol.*, 1955, **39**, 485-502.
3. PRESSEY, S. L. A group scale for investigating the emotions. *J. Abn. & Soc. Psychol.*, 1921, **16**, 55-64.
4. PRESSEY, S. L., & PRESSEY, L. C. Development of the interest-attitude tests. *J. Appl. Psychol.*, 1933, **17**, 227-38.

Moosehaven Research Laboratory
Orange Park, Florida

THE EFFECT OF ORGANIZATION UPON COMPLEX REACTION TIME*

Nuclear Development Corporation of America

IVAN FLORES

A. AIM

The purpose of this experiment was to show that organization of either the stimulus or the response field, or both, of a complex situation resulted in a reduction in reaction time to the situation.

A literature search revealed that it has been taken for granted that organization has a beneficial effect upon reaction time to a complex situation.

1. Definitions

A few definitions are required to simplify the exposition.

a. Complex reaction time. The subject is in the test situation where stimulation to only one sense organ is changed. The change of stimulation may take place in one of the several forms that the stimuli assume (here, several numbered lights). This change is of the same quantity in each form (a light is either on or off). For each form there is a required response. The quantity of stimulation of one of these forms is now changed. The elapsed time between the initiation of the change and the subject's response, provided it is the prescribed response, and his first completed response, is the complex reaction time. In case some other response than the prescribed response is completed by the subject, this is considered as an error response. For the purposes of this study, responses to stimulation of more than one sense are considered to be of a higher order of complexity.

b. Linear organization. Linear organization of either stimulus or response units will be defined as an arrangement of the units in a straight line perpendicular to the line of sight of the subject and with the units in order of increasing magnitude of their numbers from left to right.

c. Increasing organization. For this experiment, "increasing organization" or "greater organization" means changing the organization of the fields from an earlier member to a later member of the following series: randomly organized stimulus and response fields; randomly organized response field

*Received in the Editorial Office on November 26, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

and linearly organized stimulus field; randomly organized stimulus field and linearly organized response field; linearly organized stimulus and response fields.

d. Organization of the field. If the stimulus or response objects are placed in their field so that they follow no set pattern, the field will be termed not organized, or alternately, randomly organized. If, on the other hand, knowing the position of all but one of the objects, there exists a formula of degree less than the number of objects on hand, which determines the position of the last object, the field is said to be organized.

e. Degree of complexity. The number of forms of the stimulus and response variables which are subject to change in the situation is known as the degree of complexity of the situation. In this experiment there are 10 lights which might be illuminated, hence the degree of complexity is 10.

2. Assumptions

Complex reaction time increases with the degree of complexity of the situation. This was established by Hick (1) in 1952.

Complex reaction time is a linear function of simple reaction time and of some function of the degree of complexity of the situation and of only these two items. Hence, if a factor does not affect simple reaction time, it does not affect complex reaction time. Similarly, if a factor affects simple reaction time in some systematic manner, it will affect complex reaction time in the same manner.

Hick (2) displayed to each of his subjects a set of lights arranged in a circle. To each light corresponded a button, on each of which one of the subject's fingers rested. On different runs, a different number of lights was presented. This number ranged from one to 10. For each situation the subject was asked to respond by pressing the corresponding button. He found that, as the number, n , of lights involved (degree of complexity of the situations as defined above) increased, the complex reaction time increased. He plotted curves of reaction time versus degree of complexity. Similar data was obtained by the experiments of Merkel in 1885. From this data Hick worked out a formula to express the relationship between the variables. It is given here:

$$T = K \log(n + 1),$$

where T is complex reaction time, n is the complexity of the situation, K is a constant.

This formula was obtained by the combined use of empirical analysis and the current concepts of information theory. The constant factor, K ,

is a measure of the subject's simple reaction time in the environment of the experiment.

Hick's work indicates that in a fixed environment the only parameters which determine the individual's complex reaction time are the individual's simple reaction time to the sensation under study and the degree of complexity of the situation. If there were other factors, these would appear in the formula as other constants. They could not be included in K , for then this would change the value for the simple reaction time. It was felt, hence, that factors which do not affect simple reaction time will not affect complex reaction time, and those that do affect simple reaction time will affect complex reaction time in the same manner.

B. THE EXPERIMENT

1. Apparatus

Two fields were required on which to arrange the stimulus and the response media. These stimulus and response media were so designed that they could be placed as desired on punched masonite boards. Two such boards were hinged together but held rigid so that the stimulus field was vertical and the response field was horizontal. The area was allotted so that the area used fell within a visual angle of 20 degrees as suggested by the experiments by Poffenberger (3). The boards were painted flat black for no reflections and to contrast with the display units.

A diagram of the construction of the display units appears as Figure 1. There were 10 of each type for stimulus and response. Both types used identical translucent plastic squares for display. The units of each type were numbered from one to 10 with Artype opaque black letters of the same font and size for both kinds of units. A seven-watt lamp illuminated the numbered square so that the number was distinguishable many yards away. The response square was pivoted so that a few ounces of pressure and movement of one-eighth of an inch would cause the action to register.

The stimulus and response units were placed on their respective boards in one of the arrangements to be tested. They were hooked up to the presentation unit. The experimenter selected one of 25 possible random sequences of presentations for illumination of the lamps. He pressed the start button. One lamp was illuminated. The subject responded by pressing what he thought was the plastic square correspondingly numbered to the square in the stimulus field, which was illuminated. Three seconds after he had pressed a square the next lamp in the chosen sequence was displayed. This continued until 25 lamps had been displayed and reacted to. The unit

stopped. The time elapsed between each display and the subject's response was totalized and appeared on the dial of a timer where it was read in hundredths of a second.

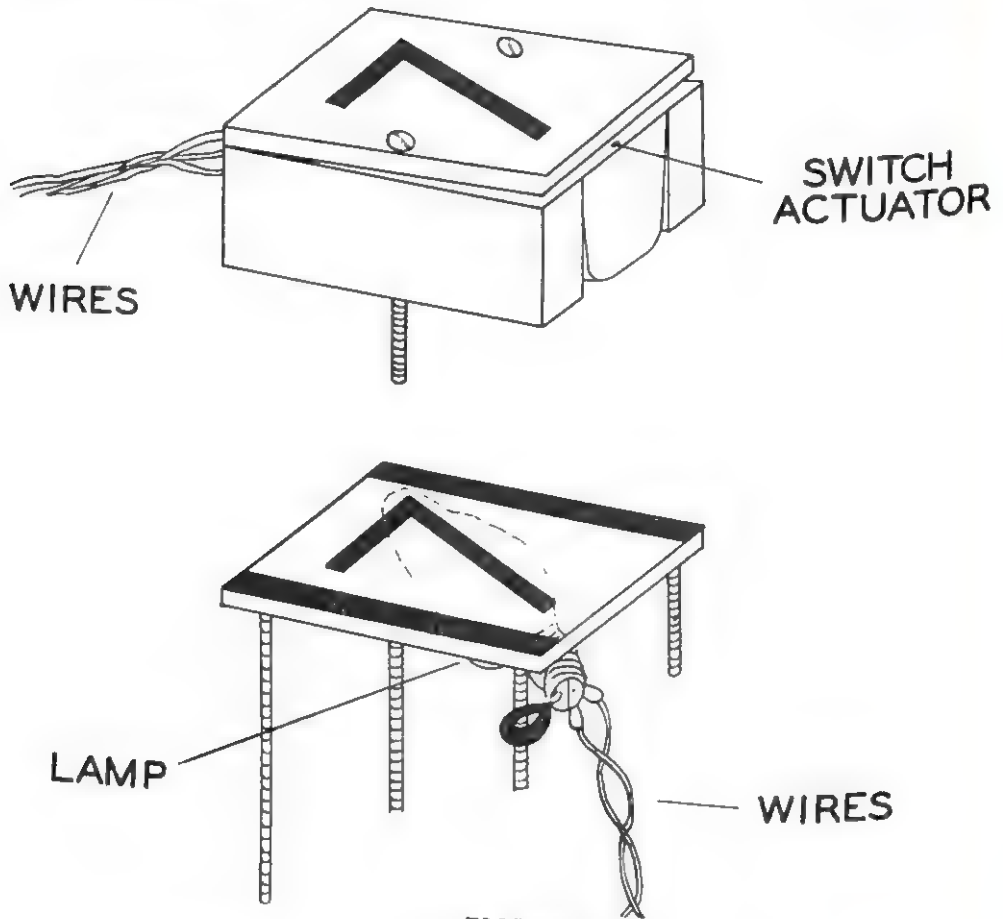


FIGURE 1
LIGHTS AND BUTTONS

When the subject pressed a button which was not numbered the same as the lamp which was lit, a red lamp lit on the panel of the presentation unit, visible only to the experimenter. During the course of this experiment, the accumulated errors for no subject was greater than 2 per cent of the trials. An error response was therefore ignored and counted as a correct response.

Four series of runs were given each subject as follows:

- First Series: Organized stimulus and organized response fields
- Second Series: Random stimulus and organized response fields
- Third Series: Organized stimulus and random response fields
- Fourth Series: Random stimulus and random response fields.

The random arrangement was made by taking the usable area and dividing it into squares. Each square was assigned a two-digit number. A table of random numbers was entered randomly. The first two-digit number determined the square where number one lamp would be placed; the next two-digit number determined where the number two lamp would be placed, and so forth.

One such arrangement was made for the second series and one for the third; two other arrangements were made so that the fourth series would introduce the subject to an unlearned situation.

In the organized situation the units were put in a linear ordered fashion; that is, in a straight, horizontal line in increasing number from left to right. In the first series the stimulus units on the vertical board were centered upon it and placed above the response units which were centered on the horizontal board.

Each series consisted of four runs of 25 trials. Each run for a given subject throughout his part of the experiment used a different presentation sequence. The order of the sequence also varied from subject to subject.

All the factors of external environment which affect simple reaction time were taken into account and held constant in the administration of the experiment. There was a uniform lighting of moderate level, a minimum of distraction, a comfortable temperature level, an adequate warning before each presentation but not too large a wait between the presentations. The subject's favorite arm was used.

Since the same equipment, environment, and subjects were used in each series of runs, it was only the organization of the situation which changed from series to series for each subject. This minimized the effect of the individual subject in the experiment.

2. *Subjects*

Ten main subjects and one supplementary subject were used.

The subjects for the experiment consisted of volunteers from among the experimenter's friends and associates. No attempt was made to choose or eliminate the subjects on any extrinsic ground. The experimenter's background, socio-economic status, profession, and so forth interposed a selective factor toward the people with whom he associated. It was felt that the factors which were operative in this intrinsic selection were those which did not affect reaction time as indicated in the literature. Thus socio-economic status, intelligence, race, or religion would not affect the reaction time of subjects to any great extent.

Subjects of both sexes were used and no obvious difference in their results could be noted.

The age range of the subjects was from 25 to 40 years. Therefore the findings would not cover childhood or senility.

Any other factor which might affect reaction time was not known. It is very likely that the population would be random with respect to such factors, as no other means was used to select the subjects besides those already mentioned.

Personal testimony from the subjects indicated that they were neither especially tired nor worried.

Directions to each subject were standardized.

3. Procedure

The experiment was performed as above using 10 subjects. It was then observed that for a given subject and series, the times for successive runs would almost invariably decrease. This indicated that situation learning was still taking place at the end of the series. In order to better consider this aspect, an additional subject was tested using a series of 15 instead of four runs. Fifteen was used because, by the end of that many runs, the learning curves had leveled off.

C. THE DATA

An analysis of variance was applied to the results of the experiment. The outcome of this is shown in Table 1.

TABLE 1
ANALYSIS OF VARIANCE FOR DATA

Description	Degrees of freedom	Mean squares
Treatments	3	2,820
Subjects	9	14.44
Cells	39	237.4
Total	159	71.13
Treatments x subjects	27	24.81
Within cells	120	17.08

To analyze the contribution of the treatments to the results, the "*F*-test" is used. The ratio of mean squares of treatments to mean squares of cells is taken as the *F* ratio. The degrees of freedom for these are respectively three and 120, and the value of *F* is 165.0. The one per cent value of *F* is found in suitable tables to be 3.94. For an *F* larger than 3.94, the results would be due to chance one time in 100. For an *F* such as that obtained with the data of the experiment, the results could hardly be due to chance. Hence

the arrangements have an almost certain effect with respect to the subjects' reaction times as set up in the experiment.

To analyze the effect on the data of the difference among the subjects, the ratio of mean squares of subjects to mean squares of cells is used as the F ratio. The degrees of freedom for each of these is respectively nine and 120 and F is 0.847. For these parameters, the F values for five per cent significance as found in the tables is 2.75. Since the value of F obtained from the data is less than one-third as large as this value, it is apparent that the differences among the subjects are similar to those which would occur if the sample were drawn from a population which had a normal probability distribution with respect to the attributes in question. Put more succinctly, the results are not likely to be due to the choice of subjects.

Having established that the treatments cause a difference in the results, it is necessary to determine whether this is due to one, several, or all of the treatments. To find the effect of the individual treatments the unbiased estimate of the variance of the difference of the treatment means, symbolized $\hat{\sigma}_D^2$ is used. This is given by,

$$\hat{\sigma}_D^2 = \frac{2}{n} \cdot \frac{m..}{s}$$

where n is the number of measurements for each subject for each treatment, s is the number of subjects and $m..$ is the mean squares for treatments times subjects. Then,

$$\hat{\sigma}_D^2 = \frac{2}{4} \cdot \frac{24.81}{10} = 1.24 \text{ sec}^2$$

and

$$\hat{\sigma}_D = 1.11 \text{ secs}$$

where σ_D is the unbiased estimate of the standard deviation of the difference of the treatment means.

Student's t -test is used to find the level of significance of the difference of the treatment means. t for the 1 per cent level is 2.473 and for the 0.05 level is 3.690. To find the smallest difference of means of treatments which

would be significant for these levels, merely multiply these values by $\hat{\sigma}_D$ thus

$$\begin{aligned} d_{1\%} &= t_{1\%} \hat{\sigma}_D \\ &= (2.473)(1.11) \\ &= 2.74 \text{ seconds} \end{aligned}$$

$$\begin{aligned} d_{0.05\%} &= t_{0.05\%} \hat{\sigma}_D \\ &= (3.690)(1.11) \\ &= 4.09 \text{ seconds} \end{aligned}$$

In Table 2 which follows, the differences between the treatment means taken two at a time are presented. All of these are greater than the value for significance at the 0.05 per cent level, except one which is significant at the 1 per cent level. Hence all the treatments are very significantly different from each other.

TABLE 2
DIFFERENCES OF MEAN ACCUMULATED REACTION TIMES

Minuend	Values are in seconds Subtrahend		
	Organized stimulus only	Organized response only	Both stimulus and response organized
Random organization of both stimulus and response			
Organized stimulus only	2.65	12.09	19.13
Organized response only		9.44	16.48
			7.04

Note—Significant difference at 0.05 per cent level is 4.09.
Significant difference at 1.00 per cent level is 2.74.

From the table it is seen that the organization of the response field will reduce the reaction time much more than organization of the stimulus field. Organization of both fields has more effect than organization of either field alone. In fact, the reduction in reaction time thus obtained by organizing both fields is greater than the sum of the reductions from both organization of the stimulus field alone and organization of the response field alone.

4. The Effects of Learning

It would appear that sequence learning, the anticipation of what the next lamp to be lit will be, was eliminated. Enough sequences were available so that a different one was used for each run for a given subject.

Familiarity with the situation, which might give the subject a facility in reacting in the experiment, was used so that it could only hinder the purposes of the experiment. That is why the series with expected greatest reaction time was given last. If facilitation existed, it would act to shorten what was hypothesized to be the longest reaction time.

Field learning is the name applied to the process by which the subject becomes familiar with the stimulus and response fields so that he knows just where to reach to respond to the illuminated lamp. The work of O. A. Simley (4) has shown that this phenomena follows a definite law. Supplemental data was taken with a subject until it appeared that his learning had reached a saturation. These results may be found in graphical form in Figure 2. For each situation the accumulated time for each run is plotted

against its occurrence in the series for that situation for the supplemental individual. For each situation, the mean accumulated time for the sample for each run is plotted against its occurrence in the series for that situation. This has been extrapolated to correspond to the supplemental data.

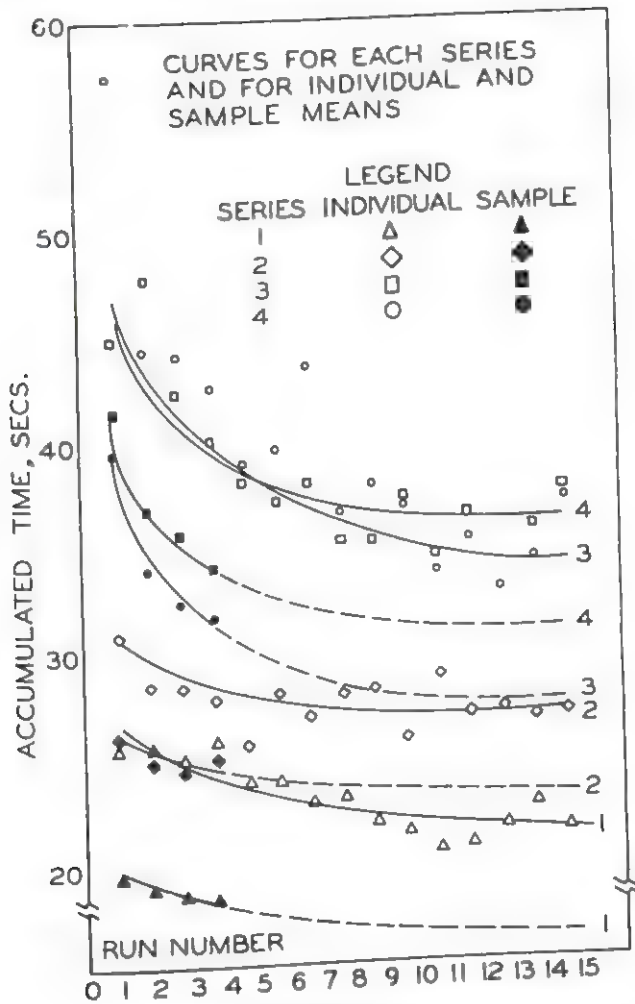


FIGURE 2
ACCUMULATED REACTION TIME FOR EACH RUN

Although the reaction time of the individual used in the supplementary data is somewhat higher than the mean of the sample, it is still useful for extrapolating the data of the sample.

A comparison has been previously made among the differences of the total

series means. Since field learning is present, there is a tendency for the reaction time of the individual to decrease as the individual's exposure to the situation becomes longer. Therefore, a more representative measure of the individual's reaction time after learning is the accumulated time for the last of the four runs comprising the series for a given situation. The mean last run for the sample would be the mean of these times taken for all the subjects. The differences among the mean last runs for the situations studied may now be taken. The differences of the mean last runs are tabulated in Table 3.

TABLE 3
DIFFERENCES OF MEAN LAST RUNS FOR EACH SERIES FOR THE SAMPLE

Minuend	<i>Values are in seconds</i>		
	Subtrahend		
	Organized stimulus only	Organized response only	Both stimulus and response organized
Random organization of both stimulus and response	2.25	8.52	15.77
Organized stimulus only		6.77	13.52
Organized response only			7.25

Note—Significant difference at the 0.05 per cent level is 4.09.
Significant difference at the 2.5 per cent level is 2.27.

To find the significance of these differences, the criteria developed in the previous chapter may be used. These criteria were based on an over-all variance of the sample. Since learning was shown to have taken place, the variance should be calculated with respect to the learning curve rather than with respect to the mean for the four runs in the series. Variance based on the learning curve would be smaller than that based on the mean. This would result in a smaller criteria for gaging the significance of the differences. This would mean a smaller difference would be significant on the same level than would result from the use of the old criteria. Hence, it is seen that the previously established criteria are more stringent than would result if they were based on variance calculated by using the learning curve.

Applying the criteria to the differences it is obvious that all the differences are significant at the 0.05 per cent level except that between the third and fourth series which is significant at the 2.5 per cent level.

5. Conclusions

The following conclusions are based upon the analysis of the statistical data and are at a level of significance of 0.05 per cent, except for the statement with the asterisk, whose level of significance is 1 per cent. These conclu-

sions presume that little or no field learning took place before exposure to the situation and that the exposure lasts for a limited length of time.

1. The complex reaction time to a situation with randomly organized stimulus and response fields is greater than the complex reaction time to a situation with linearly ordered stimulus and randomly organized response field of the same degree of complexity.

2. The complex reaction time to a situation with randomly organized stimulus and response fields is greater than the complex reaction time to a situation with randomly organized stimulus field and linearly ordered response field of the same degree of complexity.

3. The decrease in complex reaction time, compared with a situation with randomly organized stimulus and response fields effected by a linear ordering of both stimulus and response fields, is greater than the sum of (a) the decreases effected by a linear ordering of the stimulus field while the responses field is randomly organized, and (b) the decrease effected by the linear ordering of the response field while the stimulus field is randomly organized, the degree of complexity of the situation remaining constant.

Upon the consideration of the evidence brought forth, the above statements may also be made for situations in which learning has proceeded to its physiological limit. This is based on (a) the assumption of a systematic learning process as postulated in the literature, (b) a comparison of the learning curves of an individual with the learning curves for the sample and (c) a statistical analysis of the trend of the data, that is, of the mean last accumulated reaction for each series. The statistical analysis for this phase is significant at the 0.05 per cent level for the statement followed by the asterisk, for which the level of significance is 2.5 per cent.

Extension to situations of other degrees of complexity follows from Hick's work, previously quoted, on the relation between reaction time and the degree of complexity of the situation. A limitation arises when the degree of complexity becomes so small that no arrangement can be said to be randomly organized, two or three degrees, perhaps.

D. DISCUSSION

The conclusions of the experiment show that (a) if the operator is new to the situation, organizing the stimulus field will facilitate his activity; (b) if the operator will thoroughly learn the situation, organization of the stimulus field alone will have little or no effect upon the speed of his reactions.

This result follows from the kind of stimulus situation studied; any change in the stimulus field immediately alerts the operator. The change

itself informs the operator of the new information existing in the situation. Also one and only one change takes place at one time.

Since the information is immediately available, it might be only the response situation which the operator must learn. Once the response situation is known thoroughly to the operator, it is no longer of much importance how the data are presented to him.

The arrangement of the response media is of prime importance. The time required to make the judgments: "Do I make a reaction?" and "What reaction do I make?" seems small compared to the time required for the more complicated non-verbalized question: "How do I make the proper response?" This requires the unconscious selection of an intricate combination of muscles. If the reactions specified call for similar muscle groups *and* if the method of determining which muscles are required for which response, hinged upon a simple law, then it would seem that the response would be facilitated. This is the case when organization is imposed upon the response field.

When the same organization is imposed upon both stimulus and response fields a much greater facilitation results than if only the response field were organized. This might be explained in terms of the cueing by the stimulus and the stimulus field of certain organizational and muscular processes which militate for the resulting facilitation. Thus the stimulus field might help circumvent some of the time-consuming efforts of the individual in reacting to the stimulus. Also the eye and the attention have been drawn to the stimulus unit by its activation. If the response unit is in the same small, one-degree angle of vision of small detail used to observe the stimulus unit it would seem easier to find it. Otherwise, the eye would have to move, refixate, and then muscular directions would have to be given to move the arm to the proper response unit.

Although it had been previously tacitly assumed that organization of a situation led to a reduction in the reaction time to it, there had been no demonstration of where organization could best be applied. More important, the great effectiveness of organization of both stimulus and response fields in a congruent manner may now be recognized and practical designs revised to incorporate this principle.

E. SUMMARY

In this experiment, two fields, each of 10 squares numbered 1 through 10 were presented to the subject. The subject responded by pressing a square in the response field corresponding to the one in the stimulus field which had become lit. Four runs of 25 stimuli were presented to each of 10 subjects

for each of four types of organizations. These latter were: linearly organized stimulus and response fields; linearly organized stimulus and randomly organized response fields; randomly organized stimulus and linearly organized response fields; randomly organized stimulus and response fields. The experiment was performed automatically so that a constant period of three seconds occurred between the subject's response and the subsequent presentation.

Analysis of variance as applied to the results of the experiment showed differences in treatments significant at the 0.05 per cent level, except one significant at the 1 per cent level.

Learning was shown to occur and when its effects were eliminated the difference in treatments was still significant at the 0.05 per cent level, except one significant at the 2.5 per cent level. This was when learning had proceeded to its physiological limit.

It was further demonstrated that a greater reduction in reaction time to a complex situation with randomly organized fields was achieved by organization of both fields than the sum of the reductions achieved when either field was separately organized.

REFERENCES

1. HICK, W. E. On the rate of gain of information. *Quart. J. Exp. Psychol.*, 1952, 4, 11-26.
2. ———. Information Theory in Psychology. *Trans. Inst. Radio Engs.*, Professional Group on Information Theory, PGIT-1, February 1953. (Pp. 130-134.)
3. POFFENBERGER, A. T. Reaction time to retinal stimulation. *Arch. of Psychol.*, 1912, No. 23.
4. SIMLEY, O. A. The relation of subliminal to supraliminal learning. *Arch. of Psychol.*, 1933, No. 146.

Nuclear Development Corporation of America

5 New Street

White Plains, New York

A STUDY OF AUTHORITARIANISM AND PSYCHOPATHOLOGY*

Mary Conover Mellon Foundation, Vassar College

MERVIN FREEDMAN, HAROLD WEBSTER, AND NEVITT SANFORD¹

A. PROBLEM

The Authoritarian Personality (1) left the question of the relation between psychopathology and authoritarianism or ethnocentrism quite up in the air. Thus, Maria Levinson, in concluding her special study of 121 clinic patients, writes: "Our investigation shows that one is likely to find people with more or less severe psychological disturbances in the high, low, and middle quartiles [of the distribution of scores on the Ethnocentrism scale] although we cannot say in what proportions. It even suggests the possibility that most disturbed people will be found in the middle quartiles" (1, p. 968).

In this light it seems a little odd that one should encounter so frequently the notion that *The Authoritarian Personality* regarded high scorers on authoritarianism as mentally unhealthy, low scorers as healthy. Masling (4), for example, expresses this view and then reports that his results show *no* relations between authoritarianism and certain measures of psychopathology. More recently Almond (2) has written: "Shils has already stressed the Marxist presuppositions of the study which led its authors to associate right-wing extremism with psychological disorders, while the left tended to come off with a clean bill of mental health."

A possible explanation of these misinterpretations is that writers such as Masling and Almond noted—or encountered other people who had noted—the California investigators' greater liking and approval of their relatively unprejudiced subjects and assumed that they attributed everything good, including mental health, to them. Our reading of *The Authoritarian Personality* tells us that its authors liked the low authoritarians despite, or perhaps because of, the signs of mental ill health that were common among them, and that their main contribution to the problem before us was a number of hypotheses, and some suggestive findings, concerning differences between high and low authoritarians in susceptibility to different kinds of mental disturbance.

*Received in the Editorial Office on December 5, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹This is an expanded version of a paper entitled, "Some psychodynamic correlates of authoritarianism in women," presented at the sixty-third annual convention of the American Psychological Association, San Francisco, September, 1955.

These hypotheses and suggestions have not heretofore been followed up. They concern the psychodynamics of authoritarianism, a matter which, though it was the focus of the research leading to the publication of *The Authoritarian Personality*, has been rather neglected in the studies and discussion stimulated by that volume. In order to understand the relations of authoritarianism and mental health-ill health it is necessary to go at least as deep into psychodynamics as did the authors of the original study and with the use of adequate sampling. The present study takes a step in this direction; it relates responses on the *MMPI*, by large groups of young women, to scores on the *F* scale for measuring authoritarianism.

B. RESEARCH PROCEDURES

In two samples² of Freshmen, $N_1 = 441$, $N_2 = 225$, the relationships between the *MMPI* and *F* were investigated in three ways as follows: (a) *F* scores were correlated with the standard clinical *MMPI* scales (Table 1). (b) Each *MMPI* profile was categorized on the basis of its most prominent clinical feature, e.g., compulsion neurosis, hysterical tendency, schizoid trends (12 such categories in all) and the mean *F* scores for these groupings compared (Table 2). (c) The *MMPI* categorizations described above were studied in relation to three groupings based on *F* score: high, middle, and low (Table 3).

No attempt was made to classify *MMPI* profiles in which all the scores were lower than 60. These constitute about 8 per cent of each sample and appear as the < 60 category of Table 2. Otherwise, all profiles were classified. Table 2 contains a brief account of the prominent features of each category.³ Needless to say the *MMPI* categorization was carried out without knowledge of the *F* scores of the subjects involved. Also, for purposes of studying the relationship between elevation of the *MMPI* profile and *F*, the *MMPI*'s not falling into the < 60 category were divided into two groups, ≥ 60 but < 70 , and ≥ 75 ; the former indicates a profile with no scores greater than 69 and at least one equal to or above 60, the latter indicates at least one score equal to or exceeding 75.

C. RESULTS

The outstanding feature of Table 1 is the correlation, $-.41$ in the first sample, $-.44$ in the second, between *F* and the *MMPI* Hysteria Scale

²The 225 subjects were a random sample of a total class, $N = 369$. They were selected by using a table of random numbers.

³One rater was considered sufficient for sorting purposes. Previous studies in which the senior author has participated have demonstrated that where the criteria for categorization are spelled out in detail, inter-rater agreement is high.

(minus its somatic items). Typical items of this scale are ones stressing friendliness and social ease, confidence in human nature, and a rather generalized denial of aggressive or evil tendencies in the self or in others; in short, impunitive modes of handling hostility are stressed. Examples of such items, with the direction of response of low authoritarians are:

I think most people would lie to get ahead (False).

I like to read newspaper articles on crime (False).

At times I feel like swearing (False).

I can be friendly with people who do things which I consider wrong (True).

TABLE 1
CORRELATIONS OF *F* SCORES WITH *MMPI* CLINICAL SCALES

<i>MMPI</i> Scale	Sample 1 <i>N</i> = 441	Sample 2 <i>N</i> = 225
L	-.04	.03
F	.00	.00
K	-.18*	-.30*
Hs	.10*	-.02
D	-.05	.00
Hy	-.02	-.03
Hy minus somatic items	-.41*	-.44*
Pd	.03	-.09
Pa	-.16*	-.54*†
Pt	.13*	-.02
Sc	.07	-.01
Ma	.05	-.15*

*Significant at the 5 per cent level.

†*N* = 369.

TABLE 2
MEANS AND STANDARD DEVIATIONS OF *F* SCORES FOR *MMPI* DIAGNOSTIC CATEGORIES

<i>MMPI</i> Grouping*	Sample 1 <i>N</i> = 441			Sample 2 <i>N</i> = 369		
	<i>N</i>	\bar{x}	<i>s</i>	<i>N</i>	\bar{x}	<i>s</i>
Repressive—Primary elevations on <i>L</i> , <i>K</i> , and <i>Hy</i>	134	110.81	25.50	106	113.92	24.00
Compulsive—Primary elevation on <i>Pt</i> , sometimes accompanied by elevations on <i>D</i> and <i>Sc</i>	25	123.59	23.92	30	123.03	20.89
Schizoid—Primary elevations on <i>Sc</i>	29	115.14	17.88	15	108.00	21.90
Psychopathic—Primary elevation on <i>Pd</i>	17	111.05	21.02	18	104.82	31.11
Hypomanic—Primary elevation on <i>Ma</i>	53	118.23	33.38	55	113.58	21.42
< 60	37	119.91	22.17	34	114.11	20.36
60 < 70	272	115.15	26.58	222	114.53	19.61
75	40	113.70	24.11	42	114.19	27.59

*Only those groupings are listed which comprise 5 per cent or more of the total in at least one sample.

This finding is buttressed by the results of other types of analysis of the data. Examination of Table 2 shows that the mean *F* score for the repressive or hysterical grouping is the lowest in Sample 1, and although several are lower in Sample 2, none is significantly so at the 5 per cent level. (In both samples the mean of the repressive groups is significantly lower than that of the compulsive category. This will be discussed below.) Similarly for Table 3, for the first sample there are 51 repressive *MMPI*'s in the High *F* group, 58 in the Middle, and 69 in the Low. The difference between the high and low group frequencies is significant at the 2 per cent level. A similar trend appears in the second sample, although the results are not significant at the 5 per cent level.

TABLE 3
RELATIONSHIPS BETWEEN *MMPI* DIAGNOSTIC CATEGORIES AND *F* GROUPS

<i>MMPI</i> diagnostic category*	High <i>F</i>		Middle <i>F</i>		Low <i>F</i>	
	Sample 1 <i>N</i> = 147	Sample 2 <i>N</i> = 123	Sample 1 <i>N</i> = 146	Sample 2 <i>N</i> = 123	Sample 1 <i>N</i> = 148	Sample 2 <i>N</i> = 123
Repressive	51	38	58	43	69	43
Compulsive	16	15	9	15	6	7
Schizoid	8	6	14	9	12	8
Psychopathic	6	4	5	7	6	6
Hypomanic	21	17	18	20	16	18

*See Table 2 for description of categories. Only those groupings are listed which comprise 5 per cent or more of the total in at least one sample.

In order to investigate the possibility that the relationship between the Hysteria Scale (minus its somatic items) and *F* might differ for high *Hy* scores as compared with those in the middle range, we studied as a separate group those *MMPI*'s which revealed rather strong repressive tendencies, i.e., ones in which one or more of the *L*, *K*, or Hysteria Scales equalled or exceeded 70, two standard deviations above the mean. Results of this analysis demonstrate that the relationship is linear. The mean *F* score of the group with strong repressive tendencies is 101.55 in the first sample (*N* = 34) and 115.19 in the second (*N* = 27) indicating that there is no tendency for *F* scores to rise with strong repressive tendency.

The negative relationship between the Hysteria Scale and *F* is supported by the correlations between *F* and the *K* and Paranoia Scales of the *MMPI*. The *K* Scale, developed as a measure of defensiveness in the test taking situation, is usually correlated with the non-somatic items of the Hysteria Scale. As Table 1 indicates, the correlation between *F* and *K* is $-.18$ in the first sample, $-.30$ in the second, both correlations being significant at

the .001 level. The Paranoia Scale correlates $-.16$ with F in the first sample, $-.54$ in the second,⁴ both correlations being significant at the .001 level.

The findings with respect to the Paranoia Scale do not indicate a positive relationship between low authoritarianism and paranoid tendency as might at first seem to be the case. The rather complicated nature of the Paranoia Scale must be considered. Actually the scale contains three distinguishable types of items: (a) Actual persecutory ideas diagnostic of projective tendency; (b) a group of items which have been termed the "Poignancy" subscale, e.g., "I think that I feel more intensely than most people do"; and (c) a group of items which have been called the "Naivete" subscale, e.g., "I tend to be on my guard with people who are somewhat more friendly than I expected." The latter two subscales are often correlated positively with the Hysteria Scale and negatively with the Persecutory Ideas subscale.

It is to these latter two subscales of the Paranoia Scale, Poignancy and Naivete, that one would look in attempting to account for the significant negative relationship between the Paranoia Scale and authoritarianism. To examine further the quality of this relationship, the correlations between the subscales of the Paranoia Scale and authoritarianism were obtained for the second sample. These were as follows: Persecutory Ideas, .01, Poignancy, $-.07$, Naivete, $-.22$. The last correlation is significant at the .001 level. As anticipated, the negative relationship between the Paranoia Scale and authoritarianism is seen to be based not on any relationship between persecutory ideas and authoritarianism but chiefly on the items comprising the Naivete Subscale which emphasize impunitive ways of handling hostility.

The degree of relationship, $-.54$, in the second sample is puzzling, however. It seems extreme and is probably the result of some idiosyncrasy of the sample which we were unable to discover. We have examined the relationships between the *MMPI* and F for other groups of women and in no other sample is the relationship between the Paranoia Scale and F so strong.

Our major finding then is of a weak but definite negative relationship between hysterical or repressive tendency and authoritarianism. Next in importance is the discovery of a rather slight but fairly consistent positive relationship between compulsion neurosis or intropunitive tendency and F . Table 1 shows that for the first sample the correlation between the Psychasthenia Scale of the *MMPI* and F is .13, significant at the 1 per cent

⁴This correlation is based on an N of 369, the total class. The random sample of 225 was extended to include all cases because of the unexpected magnitude of the original correlation between F and the Paranoia Scale, $-.68$. Including the additional 144 cases lowered the correlation from $-.68$ to $-.54$.

level. This finding does not hold up in the second sample, however, the correlation being equal to $-.02$. It is not in the correlations between the Psychasthenia Scale and F but in the other types of analysis, however, that the relationship appears most clear.

Thus, when one compares in Table 2 the mean F score of the repressive or hysterical *MMPI* grouping with that of the psychasthenic or compulsive category, the mean of the repressive group is lower—and significantly so in both samples—at the 2 per cent level in the first sample, at the 5 per cent level in the second. Similarly, when we compare in Table 3 the number of compulsive *MMPI*'s in the high, middle, and low parts of the F distribution, we find a tendency for the number in the high group to exceed that in the low. For the first sample the number in the high group exceeds that of the middle at the 10 per cent level of significance, that of the low group at the 1 per cent level. In the second sample, we do not find exactly this relationship, the number of compulsive *MMPI* profiles in the high and middle groups being equal, but the difference between the middle and low groups reaches the 5 per cent level.

In short there seems to be a small but consistent positive relationship between F and intropunitive tendency as measured chiefly by the compulsive category of the *MMPI* groupings. This category as indicated in Table 2 is distinguished chiefly by elevation of the Psychasthenia Scale. The items of this scale stress conative difficulties, guilt feelings, feelings of anxiety and inferiority, as well as actual compulsive tendencies.

The remaining findings are of a negative nature. According to the data offered in Tables 1, 2, and 3, F is not related to elevation of the profile, i.e., to degree of over-all psychopathological tendency as measured by the *MMPI*. Moreover, it is not related to schizoid or schizophrenic tendency nor to extrapunitive tendency as measured by the Psychopathic Deviate and Mania Scales.

D. DISCUSSION

It may be noted here that our subjects average slightly below 18 years of age, are highly intelligent, are primarily of upper-middle class status, and as a group display little psychopathology. As mentioned above, our major findings have been borne out in other samples of women, including a class of seniors and some groups of college alumnae of various ages, indicating that the results are not a function of age. Caution should, of course, be exercised in generalizing the findings to men.

The importance of anal reaction formations in the character structure of authoritarian women is stressed throughout *The Authoritarian Personality*, and

our findings of a positive relationship between compulsive tendency and *F* would appear to support this. In this connection we may note that a study by Farber (3) demonstrates positive relationships between anal character traits and preference for "violent, power-oriented solutions to . . . current political problems."

Our findings may help to clarify what is an area of some confusion in *The Authoritarian Personality*. This is the matter of the relationships between authoritarianism and the various forms of handling aggression, i.e., intro-, extra-, and impunitiveness. In this book impunitiveness is sometimes treated as a quality of those high in authoritarian tendency, sometimes a quality of those who are low, while for the most part intro-punitiveness is associated with low scores. For example, the authors write: "It may well be that while they [low scorers] succeed more often than do high scorers in avoiding manifestations of aggression which are destructive of others, they do so at the price of increased self-destruction" (1, p. 451). However, on occasion attention is called to attitudes of underlying self-contempt and the like which *high* authoritarians possess. Lastly extrapunitiveness is associated with high authoritarianism.

Our findings, while not strictly comparable to those of *The Authoritarian Personality*, since the latter are based primarily on interview data, suggest the following. Extrapunitiveness, in the sense of relative freedom in the expression of aggressive or dominating tendencies, as measured by the Psychopathic Deviate and Mania Scales, is unrelated to authoritarianism. Impunitiveness, as measured by the Hysteria Scale (minus its somatic items), is negatively related to authoritarian tendency; while intro-punitiveness, in the sense of guilt feelings and attitudes of inferiority, as found in the items of the compulsive category of the *MMPI* profile groupings, is positively related.

It may be that the apparent discrepancy between our findings and those reported in the original study is a function of semantic difficulty or confusion of levels of functioning. Thus, intro-punitiveness in the sense of some sort of "true" conscience which permits realistic self-criticism and appraisal would go with low authoritarianism on theoretical grounds. Intro-punitiveness, however, in some form of early incorporation of hostility which has then been directed against the self, logically appears to be a quality of high authoritarianism. This is what one would expect to be an outcome of the kind of punitive, restrictive early family environment which is often characteristic of authoritarian individuals. A study by Omwake (5) supports this notion of the closeness of relationship between projection and introjection. On the basis of data obtained by the use of three personality inventories she concludes

that those who accept themselves tend to be accepting of others, while those who reject themselves hold a correspondingly low opinion of others.

E. SUMMARY

In two samples of college women relationships between the *MMPI* and authoritarianism were investigated by three different procedures. The chief finding is of a negative relationship, correlations of $-.41$ and $-.44$, between authoritarianism and the Hysteria Scale of the *MMPI* (minus its somatic items). Impunitive ways of handling hostility are stressed by this scale. Intro-punitive tendency, as measured by the compulsive category of the *MMPI* profile sorts, is found to be positively related to authoritarianism. The relationship is small but consistent. Extrapunitiveness and schizoid trends are not related to authoritarianism nor is degree of psychopathological tendency as measured by *MMPI* elevation. Discussion of these results is centered upon their relations to methods of handling aggressive tendencies, i.e., extra-, intro-, and impuniteness.

REFERENCES

1. ADORNO, T., FRENKEL-BRUNSWIK, E., LEVINSON, D., & SANFORD, R. N. *The Authoritarian Personality*. New York: Harper, 1950.
2. ALMOND, G. A. The appeals of communism and fascism. Paper read at a meeting of the American Political Science Association, Boulder, Colo., 1955.
3. FARBER, L. Psychoanalytic hypotheses in the study of war. *J. Soc. Issues*, 1955, 11, 29-35.
4. MASLING, J. M. How neurotic is the authoritarian? (An Abstract). *Amer. Psychol.*, 1953, 8, 402-403.
5. OMWAKE, K. T. The relation between acceptance of self and acceptance of others shown by three personality inventories. *J. Consult. Psychol.*, 1954, 18, 443-446.

Mellon Foundation
Vassar College
Poughkeepsie, New York

IS MENTAL ILLNESS MENTAL?*

School of Philosophy, University of Southern California

GEORGE WATSON

A. THE PROBLEM

Current theory on the nature of human personality appears to assume that human beings have minds that are real entities, which work according to their own laws, independently of bodily processes. Among the best known of such theories is that of Freud. Freud originated the concept of the unconscious mind as exhibiting forces and counter-forces of which the conscious mind is largely unaware, but which actually govern the principal content of consciousness. Freudian theory maintains that under certain conditions these unconscious forces and counter-forces are actually the *cause* of the psychoneuroses and functional psychoses. A person suffering from a functional mental illness, according to Freud and the psychoanalytical school, is one whose unconscious mind is "out of order"; that is, the unconscious forces and counter-forces are in "conflict." In this schema, therapy consists of "effecting a synthesis" of the components of the unconscious (3) and such therapy is entirely psychological.

In contrast to psychogenic theories of mental illness such as psychoanalysis, which treat the mind as a practically independent entity, there is a growing amount of experimental evidence which indicates the possibility that some of the psychoneuroses and the functional psychoses are in part biochemical disorders of the body, and not illnesses of the mind as such. The present article presents a review of some of this experimental evidence, together with a discussion of some of its implications for theoretical psychology.

B. THE EVIDENCE

It is known that states which are psychologically diagnosed as psychoneuroses and functional psychoses can be induced physically, as in the cases of specific avitaminoses and experimental starvation. An extensive bibliography and a summary of the current status of vitamin deficiencies in nervous and mental disease has been prepared by Peterman and Goodhart (6). The following are ex-

*Received in the Editorial Office on December 12, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹Supported by a grant from the William Clarence Kalash Fund for Research in Basic Science.

amples of the range of mental disorders induced by insufficient intakes of specific vitamins: lack of thiamine results in ideas of persecution, mental confusion, and loss of memory; lack of riboflavin causes depression, visual disturbances, disorderly thinking, inability to concentrate or perform mental work, and forgetfulness; lack of niacin causes psychasthenia, depression, anxiety, irritability, loss of memory, mania, hallucinations, and dementia; lack of pyridoxine causes epileptiform convulsions, general irritability and weakness; in addition, the lack of cyanocobalamin, biotin, and ascorbic acid are reported to cause similar types of mental disorders.

General starvation as a cause of mental illness may be illustrated by the study on starvation of the Laboratory of Physiological Hygiene, School of Public Health, University of Minnesota (4).

In the Minnesota experiment, a test group was maintained on a semi-starvation diet for 120 days. Psychological as well as physiological variables were investigated; some of the techniques employed were Rosensweig's Picture-Frustration Test, the Rorschach, the Minnesota Multiphasic Personality Inventory (*MMPI*), a free-association test, and the subject's own self-ratings.

The Minnesota Multiphasic Personality Inventory appeared to be the most sensitive of these diagnostic instruments, showing a marked elevation in the mean scores of the three psychoneurotic scales, Hypochondriasis (*Hs*), Depression (*D*), and Hysteria (*Hy*), in subjects under nutritional stress. On one of these scales, Depression, the mean score invaded the region of abnormality, judged in reference to the distribution of scores in the general population, while another mean score, Hysteria, reached the borderline of abnormality (4, pp. 868-9).

In order to determine which test items contributed to the elevation of these psychoneurotic scores, an analysis of test responses was made. In part this analysis was directed to answering the question of whether the scores obtained from persons under nutritional stress could be interpreted in the customary psychiatric manner. Or stated in other words, the question was raised whether or not nutritional stress may result in psychoneuroses.

From the *diagnostic* point of view, the answer appeared to be that in the main, "the same items contributed to the elevation of the scores on the psychoneurotic scales in the semi-starved subjects and in patients with psychiatrically diagnosed hypochondriasis, depression and hysteria but with no physical pathology" (4, p. 870). In other words, *there were no basic differences between the responses given by individuals exposed to severe nutritional stress and by patients with a clinical diagnosis of psychoneurosis.*

"Experimental neurosis," therefore, apparently can be induced entirely by nutritional means (4, p. 871).

In the Minnesota experiment, nutritional stress was obtained in the test group by a *caloric* reduction in what otherwise would have been a normally balanced diet (4, p. 38). At the termination of the starvation period, normal personalities returned to the test group under accelerated food intake. However, since at the outset of the experiment the test group did not represent a psychoneurotic population, no inferences can be drawn concerning the efficacy of nutritional therapy in clinically diagnosed psychoneuroses and functional psychoses which are not known to be due to malnutrition.

Two distinct problems are recognizable here: (a) what are the *origins* of functional mental illness, and (b) what are *therapies* for functional mental illness. It is the first of these two questions that is the principal concern of this discussion. It is obvious that prolonged nutritional stress might be a causal factor in psychoneurosis and functional psychosis, but at the same time, nutritional therapy might be ineffective in relieving such illness, due, for example, to degenerative changes in the enzyme and co-enzyme systems resulting from such prolonged nutritional inadequacies. At the same time, however, where nutritional therapy can relieve functional mental illness, this fact is partial confirmation of the hypothesis that such illness is in part caused by nutritional stress.

Consequently, the Minnesota experiment raises a critical question concerning the relation of nutritional stress to the origins of psychiatrically diagnosed mental illness. It is generally recognized that one of the principal reasons why most types of such illness are considered to be functional, or purely psychological in origin, is that no physical pathology has to date been shown to cause these disorders. In view of our rather large knowledge of the rôles specific vitamins play in a very wide variety of mental illness (6), and in view of the demonstrated deterioration of personality in general starvation, the possibility is raised that some psychologically diagnosed psychoneuroses and functional psychoses are in part due to subclinical avitaminosis, as well as perhaps to other types of sub-clinical malnutrition. Since there is presently no method of diagnosing such sub-clinical nutritional deficiencies, the customary report of "no physical pathology" in cases of functional mental illness is not completely valid.

Within the framework of the problem of the origin of functional mental illness just outlined, Watson and Comrey (7), and Watson (8),¹ made exploratory tests of the hypothesis that some "functional" mental illness, as

¹With the help of David Rabinowitz, M.D.

diagnosed on psychiatric personality tests, is in part the result of sub-clinical, multiple vitamin and mineral deficiencies.

In the Watson-Comrey experiment, a test group was administered large quantities in specially balanced formulas of all the minerals and vitamins known or believed to be important in human nutrition. The *MMPI* was employed for diagnosis and retesting, a "total score" being computed by adding the scores from eight abnormal scales, giving an objective measure of "degree" of mental illness. The placebo group as a whole showed a median improvement of one total score point, while the experimental group showed a median score improvement of 22 total score points. This difference in favor of the experimental group was statistically significant. Judgments of clinical improvement, independent of the *MMPI*, favored the experimental group at the one per cent level of confidence (7).

In the follow-up study (8) an attempt was made to differentiate between types of subjects and differences in biochemical need as represented in differently balanced vitamin and mineral formulas, with marked clinical improvements exceeding those reported in (7).

Experiments (7) and (8) provide information of interest concerning the psychoanalytic theory that all psychoneuroses and functional psychoses are purely psychological in origin. Since in these experiments some mental illness, as diagnosed by psychiatric personality tests, was apparently relieved solely by nutritional means, one probable inference is that some of such illness is *not* purely psychological in origin, but is at least in part caused by sub-clinical nutritional deficiencies.

The foregoing brief summary of some of the experimental evidence on the relation of clinical nutritional factors to the deterioration of personality indicates two principal points: (a) some states which are psychologically diagnosed as functional mental illness may originate from nutritional deficiencies, and (b) some states which are psychologically diagnosed as functional mental illness may be relieved by appropriate nutritional therapy.

C. DISCUSSION

The fact that some "functional" mental illness can both be engendered and relieved by nutritional means is one that cannot on the surface be assimilated into current psychological theories such as psychoanalysis, which consider the mind an autonomous entity, behaving according to certain dynamic principles independently of body chemistry. The following case history² of a subject whose mental illness was relieved by nutritional means is introduced to illus-

²The customary and theoretically non-significant changes have been made in this subject's history to preserve her anonymity.

trate and compare the applications of both the psychoanalytical and biochemical explanations of the origins of functional mental illness.

A 22-year-old unmarried woman asked for experimental nutritional replacement therapy almost entirely upon the urgent demands of a friend who knew in general some of the results of previous research. The young woman herself not only had no confidence whatever in the general biochemical approach being used, but was convinced, rather, that she needed psychoanalysis.

She believed that her long-standing symptoms (about seven years) were all out of proportion emotionally to the alleged biochemical cause, and to the proposed nutritional treatment; while in addition, there was little question in her mind that she was adequately nourished in the current medical view, i.e., she was not anemic and not underweight. Further, she said that she had been under expert medical care all of her life, and had also been taking "vitamin pills" by the handful for years with apparently no effect on her emotional symptoms.

She was living at the time in the family home alone with her father. Her mother had died two years previously of a heart attack, and had been a heart patient for many years. This young woman was the youngest of three children, the other two being boys considerably older, and now living with their own families in another part of the country.

She gave the following information in the background interview: Her father was an excessively ambitious physician who spent almost no time at all at home with his wife and children. If he took a vacation it was alone and consisted of a trip to a medical or scientific conference. Her recollections of childhood and early youth consisted mainly of painful memories of family arguments, of her mother having fainting spells and heart attacks as a consequence of such arguments. This young woman said that she felt her father had indirectly caused her mother's death by his cruelty, neglect, and indifference.

She reported the following symptoms: severe depression with frequent suicidal urges; frequently she would stay in her room for days, afraid to see anyone; she developed an extreme revulsion for her father, to the point of feeling like running and screaming at his sight, and she could not bear to have him approach her or touch her. She perseverated on her mother's death, and could not dissociate this event from her feeling of revulsion for her father. On the other hand, she would occasionally be overcome with remorse about her hatred for him, and would admit to herself that she loved and admired him.

Although she was convinced that the suggested treatment was "silly," she volunteered to give it a trial. During the first month her worst symptoms began to disappear. Her depressed days became less and less frequent until they did not recur. And within three months' time she stated that she was well. Her attitude toward her father became normal, and later she married.

This case is a typical example of the kind that is found in the literature of

psychoanalysis, illustrating the theory of the unconscious origin of neurotic behavior due to repressed infantile sexuality. Here is an unmarried young woman, living alone with her father; she feels guilty about her mother's death and accuses her father; she is afraid to be near him, she both hates him and loves him, and she wants to commit suicide.

The principal thesis of psychoanalysis is that psychoneurosis occurs when a conflict is present among the dynamic centers of the self, principally at an unconscious level. Freud believed that neurosis was functional, in that it represented a compromise, so to speak, between competing tendencies of the id, ego, and super ego. Normal behavior consists of a "harmonious" adjustment of the psychic energies at the disposal of each of these dynamic centers, while abnormal behavior is *prima facie* evidence that one (or more) of these centers is "too strong."

In the case of the young woman cited above, for example, her extreme depression might be said to be due to a conflict between id demands of an infantile sexual nature, and the counterattempt by the superego to prevent these demands from reaching consciousness and from being translated into motor behavior. Consequently, the psychoanalytic therapy indicated in such a case is to resolve the unconscious conflict.

However, under experimental nutritional replacement therapy, all of the alleged "symptoms of unconscious conflict" of this subject began to leave within one month's time; they were entirely gone in three months; and at the present time three years later, she is married. *Not one word was ever said to her in the context of psychotherapy*; she was simply told the number of capsules and tablets to take each day.

If the psychoanalytic theory of the origin of psychoneuroses were completely valid, it would be impossible to help emotional illness in any other way than by psychological means, since, for example, the id, ego, and superego are psychical components of the self, not physical components.

Since it is possible to relieve some functional mental illness by biochemical means, it may be true that in some cases the success accredited to psychoanalysis as a therapy for neurosis is based upon a mistaken interpretation of the rôle analysis plays in relieving emotional symptoms, when the latter are due to nutritional stress.

The research on nutritional replacement appears to support the view that psychological stress causes nutritional stress, and that nutritional stress causes mental illness. Consequently, supportive psychotherapy might be of indirect aid in relieving nutritional stress.

The psychotherapist, by providing psychological support over a lengthy

period of time, may be instrumental in reducing the amount of psychological pressure on his patient, consequently reducing the patient's rate of nutritional displacement. With some of the psychological stress thus being buffered by the therapist, the patient's dietary habits might improve; consequently, more normal nutritional intake and less emotional output could have the net result of gradual nutritional replacement over a long period of time. If the psychotherapist can actually be instrumental in helping resolve an energy-consuming conflict, this will reduce the rate of energy output, again giving the net result of gradual nutritional replacement, providing the patient has good dietary habits.

There are two considerations to be weighed in assessing the value of psychotherapy in relation to nutritional rehabilitation.

The first is that in experiments in nutritional replacement therapy, very great psychological improvements take place without the need to "resolve unconscious conflicts" by means of psychotherapy. The above cited case of the young woman with apparently severe unconscious conflicts is a fairly typical example of the "automatic" personality adjustment that occurs when the subject overcomes sub-clinical undernourishment. Such spontaneous personality regeneration is one of the most striking results of nutritional replacement therapy. The following is a report of one subject (8), and is typical: "My likes and dislikes have changed. I like different music; I wear different clothes; my business associates say 'you are a different person.'"

The second point to be emphasized is that some types of psychotherapy, especially psychoanalysis, may be more of a source of psychological pressure than they are of relief. There are very few more painful experiences in life than the emotional turmoil one can undergo during psychoanalysis. It is said psychoanalysis is the major surgery of psychotherapy, and that in order to be helped by psychoanalysis, psychoneurotics must have "considerable stamina of personality" (5, p. 159).

Consequently, if people suffering from functional mental illness are also suffering from severe nutritional displacement, further pressure of the kind represented by psychoanalysis seems unwarranted.

The foregoing considerations return us to the long-discussed problem of why only some of us become emotionally disturbed, if functional mental illness is due principally to unconscious conflict: for all of us, according to psychoanalytic theory, have just about the same burden of unconscious conflict. We all repress infantile sex impulses as well as mature sex impulses; we all repress hostility, and learn to live with frustration. The very foundation of social living is the inhibition of unrewarded responses. If this were the only cause of mental illness, then we would all be sick.

The psychoanalysts currently answered this question in part by saying that some of us perhaps are constitutionally disposed to mental illness; we can't "mature." If we re-interpret this to mean that some persons can stand less psychological stress than others, there is good evidence to support it. However, considering functional mental illness as in part a nutritional problem probably involves several factors, those of native constitutional endowment and the adequacy of life-long diet, both as related to the amount of physical and psychological stress one undergoes.

Although an optimum normal diet most certainly protects one against undue nutritional deficiencies under ordinary life stress, for some people it apparently is not in itself sufficient. In the research on nutritional replacement, dietary histories more often than not revealed food intakes generally considered adequate. Since persons eating such diets showed very marked psychological improvements with the addition of rather potent quantities of vitamins and minerals to their food intake, the normal presumption would be that in some cases there is a greater need for vitamins and minerals than is contained in a normally balanced diet of adequate caloric intake, although there are several other alternative possible explanations.

On the hypothesis that some functional mental illness originates in part from sub-clinical malnutrition, a great variety of situations might participate in causing such psychological symptoms which have no relation whatever to emotional conflict: (a) over-work, the consistent output of energy beyond the body's ability to recover. One subject who had returned to emotional normality with nutritional replacement therapy felt so good that she took a job as a waitress; within one month, even while taking added vitamins and minerals in very potent quantities, many of her symptoms returned—extreme fatigue, weeping, depression, etc. Upon cutting down her work two-thirds, she recovered. She had been attempting a job that was apparently too great for her constitutional resources. (b) Poor dietary habits, due to many things, such as ignorance, economic factors, practices which destroy appetite such as the overuse of tobacco, alcohol, etc. An example of the nutritional ignorance of some parents is found in a recent dietary survey on about 8,000 Wisconsin school children (1), which showed that 10 to 20 per cent drank little or no milk; less than half ate sufficient citrus fruit; less than one-fourth had a daily serving of green or yellow vegetables; in addition, a very large consumption of concentrated sweets was observed. Roughly one-half of these children are laying the foundation for sub-clinical malnutrition, and possibly for functional emotional upsets in later life. (c) Situations which place great nutritional demands upon the body, such as illness or childbirth. Quite

frequently women who have just borne children become disturbed emotionally, threatening and committing suicide. These people are currently treated as though their illnesses were due to long-standing unconscious conflicts originating in infantile sex drives! (d) Poor digestion, leading a person to reject foods that give him trouble, with consequent poor nutrition. Rather than eliminating important dietary items such as fruits, vegetables, and milk, which are frequently listed by subjects as being hard to digest, such persons need enzymatic digestive aids to help them assimilate the foods they need.³ Nutritional replacement with potent quantities of vitamins and minerals can aid in regenerating certain enzyme systems of the body (6), improving digestion.

Since a great many different conditions may play a part in the origin of the emotional symptoms of sub-clinical malnutrition, it should be noted that the particular cause of a given emotional state cannot be directly inferred from a descriptive knowledge of that state. For example, severe emotional depression may result variously from lack of proper food, the use of drugs such as the barbiturates, overwork, overplay, emotional shock, emotional conflict, childbirth, severe illness, unusual need for certain substances due to hereditary constitutional factors, and so on.

Further, in cases of emotional illness resulting from sustained psychological stress which in turn causes nutritional displacement, knowledge of the *precise* traumas and conflicts involved is of little interest. It is of incidental importance only that some particular trauma or conflict occurred in the subject's past. What is of interest is the biochemical *result* of such stress, since this is what must be known if help is to be given. In the case of the young woman discussed above, it was not necessary to know the details of the stressful situations in her past in order to bring about nutritional replacement. For some reason, however, there seems to be a natural propensity in some to wish to explain "like by like," as well as to treat "like by like." For example, subjects who are just beginning to feel the effect of nutritional stress, and who are consequently becoming upset emotionally, have a strong tendency to look for the cause of their upset in the immediate area of their interpersonal relations. Instead of saying, "I am upset, what is wrong with my body chemistry" (in terms of what elements are needed to restore normal emotional balance), they are apt to say, "I am upset, what did you do to me, and why did you do it." To illustrate this point, a subject who needs a regular daily allowance of rather large quantities of a complex nutritional formula

³The preparation used in (7) and (8) contained: Oxbile extract USP 125 mg.; Mycozyme 125 mg.; Betaine HCL 100 mg.; Papain 100 mg., and Pancreatin USP 125 mg.

in order to feel just right, may miss taking his tablets for a period of time. He then begins to deteriorate emotionally, in the sense that little things begin to irritate him, and he feels that everyone is "picking" on him. He now makes the following mistake: instead of correctly assessing what is happening to himself nutritionally, he falsely attempts to interpret his condition in terms of what his wife has done that she shouldn't have, etc.

Judging clinically, the greater the degree of nutritional deficiency, the greater the tendency to withdraw socially and to relive past conflicts. Withdrawal and the tendency to perseverate on what happened yesterday are among the first signs of undue nutritional stress.

On the other hand, where a very high level of nutritional replacement is maintained, even rather great emotional shocks can apparently be absorbed with relative ease. Great nutritional displacement due to severe shocks should immediately be compensated for in added replacement. If such displacement is not compensated for, then the subject keeps thinking about the event for days, weeks, or months, until, by delayed replacement through normal diet, he slowly recovers, if he ever does.

The case history given at the beginning of this article partly illustrates some of these points. This young woman perseverated on her mother's death continually. Her life had been one of continued stress, since she had been a frequent witness to family arguments and to the heart attacks of her mother. She blamed her father for the way she felt emotionally, whereas the actual reason why she was depressed and suicidal was apparently a nutritional deficiency. Consequently we say that she was interpreting her feelings in terms of social, interpersonal relations, instead of correctly in terms of the biochemical result of such relations.

If the experimenters were to have questioned her, as Freud questioned his subjects, they could easily have formed the hypothesis that the reason why the subject was ill was because something traumatic happened in the past that was *still bothering her in the present*. In addition, the subject herself would have eagerly welcomed this interpretation and believed it! Somehow, it is far more satisfying and easy for almost everyone to accept a social interpretation of a present emotional disturbance than a nutritional interpretation. However, under nutritional replacement, this subject quickly stopped dwelling on the past entirely.

On the hypothesis that nutritional stress plays a part in the origin of functional mental illness, the psychological past has a new significance. What the past apparently means is this: If a subject is primarily concerned with past events, perseverating on "what happened" to him, this is evidence that he is

to some degree emotionally disturbed and nutritionally displaced. A person who is emotionally "healthy" is one who is not concerned with the traumas and conflicts of the past, but is one who looks ahead with optimism and confidence. In order to do this, however, he must be restored by adequate rest and nutrition each day from the physical and psychological stresses of the previous day.

D. SUMMARY

Experimental evidence indicates that some functional mental illness may be a biochemical disorder of the body, and not a disorder of the mind as such. Both psychoneuroses and functional psychoses can be caused and relieved by nutritional means. This is a fact that is not compatible with current psychological theories such as psychoanalysis, which hold that all functional mental illness is purely psychological in origin, having no physical pathology.

Since it is possible to relieve symptoms of psychological disturbance solely by nutritional replacement therapy, the question occurs whether at least some of the success attributed to psychotherapy for mental illness is not based in part upon a mistaken interpretation of the rôle psychological treatment plays in helping emotional symptoms. If it is true that sub-clinical nutritional deficiencies are involved in the origin of some functional mental illness, psychological treatment could have no direct value in such cases. However, supportive psychotherapy might be of indirect aid in the process of nutritional replacement, if it is instrumental in relieving psychological pressures which are a drain upon the physical resources of the body. However, some types of psychological treatment, such as psychoanalysis, possibly may provide more stress than support, and may consequently make an emotionally ill person more ill.⁴ Since it is true that some mentally ill persons become worse under psychoanalytic treatment, part of the reason for this may be explained by the hypothesis under discussion.

The available data on the relation of sub-clinical nutritional factors to emotional disturbance appears to support the view that *psychological stress causes nutritional displacement, and nutritional displacement causes mental illness*. In addition to psychological stress, however, many other types of causal factors may also cause nutritional displacement, such as poor dietary habits, overwork, illness, poor digestion, etc. The inappropriateness of psychotherapy as a remedy for nutritional stress due to such factors as childbirth, poor diet, or indigestion, is obvious.

⁴There appears to be an inverse correlation between recovery rate and psychotherapy: "The more psychotherapy, the smaller the recovery rate" (2, p. 322). This record might be improved if psychiatrists were to combine supportive psychotherapy with nutritional replacement therapy.

Finally, on the hypothesis that some functional mental illness is in part due to nutritional displacement, knowledge of the precise psychological traumas and conflicts of the past is of no particular interest, since it cannot tell us what is needed for physical rehabilitation.

E. CONCLUSION

Experimental psychoneurosis can be induced in mentally normal subjects entirely by nutritional means. This fact raises the possibility that some persons suffering from functional mental illness may be the victims of sub-clinical nutritional deficiencies, rather than of unconscious conflict which is purely psychological in origin. Experimental tests of this hypothesis with psychologically diagnosed mentally ill subjects resulted in statistically significant improvements. Individuals that exhibited evidences of severe psychological conflict were brought into "psychological adjustment" without recourse to psychotherapy. The evidence reviewed tends to confirm the hypothesis that some types of mental illness, which have hitherto been considered to be purely "functional" in origin, have a physical basis of a biochemical nature, and consequently are not directly due to psychological trauma or conflict.

REFERENCES

1. BILLINGTON, L. K. A look at food habits of Wisconsin children. *Wisconsin M. J.*, 1955, **54**, 312-313.
2. EYSENCK, H. J. The effects of psychotherapy: an evaluation. *J. Consult. Psychol.*, 1952, **16**, 319-324.
3. FREUD, S. *Beyond the Pleasure Principle*. New York: Liveright, 1950. Pp. 97.
4. KEYS, A., *et al.* *The Biology of Human Starvation*. Minneapolis: Univ. Minnesota Press, 1950. Pp. 1385.
5. PECK, M. W. *The Meaning of Psychoanalysis*. New York: Knopf, 1931. Pp. 273.
6. PETERMAN, R. A., & GOODHART, R. S. Current status of vitamin therapy in nervous and mental disease. *J. Clin. Nutr.*, 1954, **2**, 11-21.
7. WATSON, G., & COMREY, A. L. Nutritional replacement for mental illness. *J. of Psychol.*, 1954, **38**, 251-264.
8. WATSON, G., & RABINOWITZ, D., M.D. Experimental report in preparation.

School of Philosophy
University of Southern California
3518 University Avenue
Los Angeles 7, California

SPACE LOCALIZATION UNDER CONDITIONS OF DANGER* ¹

Department of Psychology, Clark University

SEYMOUR WAPNER, HEINZ WERNER, AND PETER E. COMALLI, JR.

A. INTRODUCTION

The effect of emotional factors on perception has been a central theme in clinical psychology for some time; it has become, in recent years, a problem of importance for the general experimental area of perception. This implies that a comprehensive understanding of perception must take cognizance of interrelations between emotional and non-emotional factors. We have recently initiated a series of studies to show how this problem can be advantageously approached by an organismic theoretical framework, in particular, the sensory-tonic field theory of perception.

The basic assumption underlying sensory-tonic theory is that perception must be considered in terms of the relationship between object stimuli and the state of the organism: perception is an experience which corresponds, not to an organismic state as such, but rather to a particular relation between organismic state and impinging stimuli. On the basis of this assumption, a number of studies have been carried out which demonstrate that the state of the organism is an important variable in perceptual organization. Organismic states have been conceived of broadly, i.e., to include changes due to both emotional and non-emotional stimuli. Accordingly, in some studies the organismic state was changed experimentally by employing a wide variety of non-emotional stimuli such as auditory, postural, labyrinthian, etc.; it has been demonstrated that manipulation of organismic state by this means significantly affects spatial localization (5, 8, 9). We have recently presented evidence that emotional stimuli, such as those brought about by a situation of danger, affect space perception in terms of psychological distance (6).

The present study attempts to clarify our understanding of the effects of emotional factors on space perception in two respects: (a) the study seeks evidence that the factor of danger influences the perception of the median plane (straight-ahead) and of verticality; (b) further, it seeks evidence in

*Received in the Editorial Office on December 30, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.
¹This investigation was supported by a grant, M-348, from the National Institute of Mental Health of the National Institute of Health, Public Health Service.

support of a postulate of sensory-tonic theory referred to as functional equivalence. By functional equivalence is meant that diverse stimuli ideally lead to identical perceptual end-products. In this particular case we attempt to show that certain emotional factors operate in an equivalent manner to non-emotional factors in their effect on space localization; in this respect, two kinds of evidence are sought: (a) analogous directional shifts of straight-ahead and apparent vertical should occur due to organismic changes evoked by both emotional and non-emotional stimuli; (b) if the two factors of this sort operate simultaneously, then the perceptual end-effect should be one of co-action.

For this purpose, three experiments were carried out: the first deals with effect of danger on the position of the straight-ahead or apparent median plane; the second with the effect of danger in combination with asymmetrical placement of an object on the position of the apparent median plane; and the third experiment deals with the effect of danger on the perception of verticality.

B. GENERAL PROCEDURES

In all three experiments the factor of danger was introduced, so-to-speak, asymmetrically; that is, the *S* stood blindfolded, on a table with a precipitous edge either on his left or on his right. The specific procedure was as follows: the *S* was blindfolded and led into a darkroom. He was then told that he would be led up to a high platform and explicitly warned not to move suddenly or he would fall off. With this warning the *S*, blindfolded, was led up four steps to a platform 31 in. above the floor. He was then assisted in side-stepping slowly until he had reached the edge of the platform either on his left or on his right. Still blindfolded, the *S*, while being supported by *E*, was instructed to put his foot over the edge. This was done to make him sharply aware of the empty space that was actually beside him. While standing next to this precipitous edge, the *S* was placed in such a way that he was confronting a stimulus figure located two meters directly in front of him. Precautions were taken to insure that the *S* did not move during the course of a trial. This was accomplished by the following device: A cardboard pointer, painted with luminescent paint, was attached to the mid-line of *S*'s forehead by means of a head band. In front of *S*, fixed to and projecting out from the back wall, was another luminiscent pointer. *E* positioned *S* so that the pointer attached to the *S* coincided with the pointer projecting from the wall. By observing both pointers *E* could, in darkness, make certain that *S* did not move.

C. EXPERIMENT 1: EFFECT OF DANGER ON PHYSICAL POSITION OF APPARENT MEDIAN PLANE

For formulating a hypothesis we venture to make the following assumptions: A precipice to one side functions as a stimulus equivalent to sensory stimuli exerting a pull to that side, which is counteracted by muscular strains in the opposite direction. In previous experiments it has been demonstrated that the apparent median plane is shifted toward the side of counteractive strains evoked under non-emotional conditions like counteraction to head turning, asymmetrically applied diffuse stimulation of the retina, etc. (9). Accordingly, it was hypothesized that the physical position of straight-ahead would shift opposite the side to which danger was located.

1. Procedure

The *S* had the task of adjusting a luminous rectangle so that it appeared straight-ahead; this task was carried out under the two conditions of danger (left, right) described above. The luminous rectangle, 40 cm. wide and 20 cm. high, was broken in the center by a vertical black strip 2.5 cm. wide; the figure was the milk glass face of a light box containing a 5-watt bulb operating at 26 volts. The figure was placed at eye level at a distance of two meters from *S* and could be moved in small steps horizontally in the fronto-parallel plane on a track. The stimulus figure thus appeared to *S* as two adjacent luminous squares separated by a black strip and suspended in space. This black strip was fixated by *S*; he was required to instruct *E* to move the figure so that this strip appeared straight-ahead.

For each test condition the fixated black strip was placed at the intersection of the median sagittal plane of *S*'s body and the frontoparallel plane.

Adjustment of the luminous figure was carried out under two conditions: (a) dangerous edge on *S*'s left; and (b) dangerous edge on *S*'s right. Twenty *S*'s, 10 men and 10 women, were tested under both experimental conditions. For half of the *S*s (5 men and 5 women) Condition *B* followed Condition *A*; for the other half the sequence was reversed. Thus, a replicated 2 x 2 Latin Square design was used with 5 male and 5 female *S*s tested in each of the two sequences. Arbitrarily, positions to the left of the objective median plane were designated by minus value; positions to the right of the objective plane were designated by plus values.

2. Results

The results are consistent with the hypothesis (Table 1). Under the condition of dangerous edge located on *S*'s left, the mean physical position of the apparent median plane is relatively to the right (+6.4 cm.) compared

with location of the dangerous edge on the *S*'s right (+.3 cm.). This difference, 6.1 cm., is statistically significant at the .05 level of confidence.

TABLE 1
EXPERIMENT 1: EFFECT OF DANGER ON THE PHYSICAL POSITION OF THE APPARENT
MEDIAN PLANE

Source of variation	<i>df</i>	<i>F</i> -tests		<i>F</i>	<i>P</i>
		Sum of squares	Mean square		
Individuals	19	2,549.50	134.18		
Sequence	1	233.77	233.77	1.78*	>.05
Sex	1	82.65	82.65	<1.00*	>.05
Individuals within sex and sequence	17	2,233.08	131.36	2.40	<.05
Order	1	2.65	2.65	<1.00	>.05
Conditions	1	371.49	371.49	6.78	<.05
Error	18	986.19	54.79		
Total	39	3,909.83			

*Tested against individuals within sex and sequence; all other sources of variation tested against error.

Means (cm.)†	
Position of Dangerous Edge	
Right	Left
+0.3	+6.4

†Positions to the right of the objective median plane are designated by "+" values.

D. EXPERIMENT 2: EFFECT OF DANGER IN INTERACTION WITH ASYMMETRICAL EXTENT ON THE PHYSICAL POSITION OF THE APPARENT MEDIAN PLANE

The next step in our analysis toward the understanding of the operation of emotional factors in regard to space localization concerns the problem of equivalence of emotional and non-emotional factors under simultaneous introduction of both factors. Thus, here, the emotional factor, danger, is systematically varied in combination with a non-emotional factor, viz., asymmetrical placement of the test object. It was found previously that the apparent median plane shifts toward the side to which the figure extends asymmetrically. In our view, this effect reflects a tendency toward symmetrization (5). The non-emotional factor of asymmetry is expected to operate co-actively with the factor of danger which singly functions to shift the apparent median plane opposite the side of danger.

Consistent with the postulate of functional equivalence and the assumption of co-active operation of emotive and non-emotive factors, the following specific hypothesis can be formulated: if the two factors—one emotive and the other non-emotive—have single effects in a particular direction, then they should show the strongest combined effects in that direction. That is

when both factors make for a shift to the right (extent right, danger left) the combined effect should result in the greatest shift of the apparent median plane to the right; when both factors make for a shift to the left (extent left, danger right) the combined effect should yield the greatest shift of the apparent median plane to the left. When the single effects are in the opposite direction, the combinatory effect should fall between these extremes.

1. Procedure

The procedure in this experiment was identical with that of the first experiment with the exception that the stimulus figure differed from that used in the previous experiment. Here, the stimulus figure was a 20 cm. luminous square placed in two asymmetrical positions: in one condition the right edge of the figure was placed in the objective median plane so that the figure extended 20 cm. asymmetrically to the *S*'s left; in the other condition, the left edge was placed in the objective median plane so that the figure extended asymmetrically to the *S*'s right. The task was, under two danger conditions, to fixate the edge placed in the objective median plane and to adjust the figure so that the fixated part appeared straight-ahead.

Thus, four test conditions were employed: (a) Asymmetrical extent left—danger left; (b) Asymmetrical extent left—danger right; (c) Asymmetrical extent right—danger left; (d) Asymmetrical extent right—danger right. A replicated 4 x 4 Latin-Square design was employed. Sixteen subjects were tested, 2 men and 2 women in each of the four sequences.

2. Results

The results are presented in Table 2. Examination of the table show that the over-all effects of danger and the over-all effects of extent are significant, but the statistical interaction of danger and extent ($D \times E$) is not significant. The effect of danger, independent of asymmetrical extent, is completely in keeping with that obtained in the previous experiment, namely, with dangerous edge on the *S*'s left the apparent median plane shifts relatively to the right (+3.6 cm.) compared with dangerous edge on the *S*'s right (—1.8 cm.). The effect of asymmetrical extent, independent of danger, is the same as obtained in earlier experiments, namely, the apparent median plane shifts in the direction to which the figure extends: the physical location of the apparent median plane is relatively to the right (+6.5 cm.) when the figure extends asymmetrically to the right and the apparent median plane is relatively to the left (—4.7 cm.) when the figure extends asymmetrically to the left.

The effect of simultaneous operation of both factors can be seen from the means for each of the four test conditions. Here the results are in keeping with the hypothesis that the two factors operate co-actively: the greatest shift of the apparent median plane to the right (+8.1 cm.) occurs when

TABLE 2
EXPERIMENT 2: EFFECT OF DANGER IN INTERACTION WITH ASYMMETRICAL EXTENT ON THE PHYSICAL POSITION OF THE APPARENT MEDIAN PLANE

Source of variation	df	F-tests			P
		Sum of squares	Mean square	F	
Individuals	15	802.44			
Sequence	3	59.19	19.73	<1.00*	>.05
Sex	1	3.06	3.06	<1.00*	>.05
Individuals within sex and sequence	11	740.19	67.29	<1.00	>.05
Order	3	129.82	43.27	<1.00	>.05
Conditions	3	2,540.82	846.94	11.40	<.01
Danger (D)	1	484.00	484.00	6.51	<.05
Extent (E)	1	1,980.25	1,980.25	26.65	<.01
D x E	1	76.57	76.57	1.03	>.05
Square uniqueness	6	499.11	83.18	1.12	>.05
Error	36	2,675.25	74.31		
Total	63	6,647.44			

*Tested against individuals within sex and sequence; all other sources of variation tested against error.

Asymmetrical extent	Means (cm.)†		Over-all mean extent
	Position of Dangerous Edge		
Right	Right	Left	
Left	+4.8	+8.1	+6.5
Over-all mean danger	-8.5	-0.8	-4.7
	-1.8	+3.6	

†Positions to the left of the objective median plane are designated by "-" values; positions to the right of the objective median plane are designated by "+" values.

both factors make for a shift to the right (extent right—danger left) and the greatest deviation of the apparent median plane to the left (-8.5 cm.) occurs when both factors singly make for a shift to the left (extent left—danger right); the two remaining conditions, where the single factors have potencies in opposite directions (extent right—danger right; extent left—danger left) fall between these extremes.²

²The problem remains whether this co-action can be viewed as additive or non-additive (configurational). The relative shifts obtained in the experiments under simultaneous operation of two factors may occur independent of whether the co-action is additive or non-additive (configurational). In a factorially designed experiment the presence of statistically significant interactions force a non-additive interpretation of co-action; however, the presence of statistically significant single effects together with non-significant interactions suggests an additive co-action. Since the single effects of danger as well as extent are statistically significant but the interaction of Danger x Extent is not statistically significant, the data are not inconsistent with an additive interpretation of co-action.

E. EXPERIMENT 3: EFFECT OF DANGER ON PERCEPTION OF VERTICALITY

Whereas in the previous two experiments the effect of danger was studied in terms of displacements of the apparent median plane, here the effect of danger was studied in terms of displacements of the apparent vertical. By apparent vertical is meant the physical position of a rod seen as vertical. One aim of this experiment, again, is to obtain further evidence on equivalence of emotional versus non-emotional factors. For this purpose, we introduce as a non-emotional factor, that of so-called "starting position." This factor is known from previous experiments (7) to affect apparent verticality: the starting position effect pertains to the fact that the apparent vertical is located relatively close to the position in which a rod is started at the beginning of a trial: e.g., if a rod is initially tilted to the left (left starting position) the final physical position in which it is perceived vertical is tilted more to the left as compared with right starting position. Analogous to the expectation for the apparent median plane, it was hypothesized, first, that the position of the apparent vertical would tilt relatively opposite the side at which the danger is located and, secondly, that the starting position factor will operate co-actively with the factor of danger in perception of verticality.

1. Procedure

The danger conditions were introduced in an identical manner to that employed in the previous two experiments; the task was to adjust a luminescent rod whose center was located directly ahead of *S* at eye level and could be rotated in the frontoparallel plane so that it appeared vertical. The luminescent rod, 39" long and 1" wide, was pivoted at its center and its angular displacement was read off in fractions of a degree from a protractor.

Two starting positions of the rod, 30° left (counterclockwise), 30° right (clockwise), were employed under each of the conditions of danger making a total of four test conditions: (a) Danger right, starting position 30° left (CCW); (b) Danger right, starting position 30° right (CW); (c) Danger left, starting position 30° left (CCW); (d) Danger left, starting position 30° right (CW). A replicated 4 x 4 Latin-Square design was employed. Sixteen *Ss* were tested under each of the four conditions, 2 men and 2 women in each of the four sequences.

The position in which the rod looked vertical to *S* (apparent vertical) was measured in degrees of tilt with reference to objective vertical (plumb line). Thus, all measures of the position of the luminiscent rod are defined in terms of the perpendicular. Arbitrarily we designate an angular position clockwise (CW) of objective vertical as "+" (tilted right) and an angular position counterclockwise (CCW) as "-" (tilted left).

2. Results

The results are presented in Table 3. Examination of the table shows that the effects of danger—-independent of starting position—are analogous to those obtained in the previous experiments, namely, with the dangerous edge on the *S*'s left the apparent vertical is tilted relatively to the right ($+1.33^\circ$)

TABLE 3
EXPERIMENT 3: EFFECT OF DANGER IN INTERACTION WITH STARTING POSITION ON THE PHYSICAL POSITION OF THE APPARENT VERTICAL

Source of variation	df	F-tests			P
		Sum of squares	Mean square	F	
Individuals	15	74.38			
Sequence	3	6.78	2.26	$<1.00^*$	$>.05$
Sex	1	5.64	5.64	1.00^*	$>.05$
Individuals within sex and sequence	11	61.96	5.63	5.41	$<.01$
Order	3	4.53	1.51	1.45	$>.05$
Conditions	3	17.66	5.89	5.66	$<.01$
Danger	1	6.89	6.89	6.62	$<.05$
Starting position (<i>S</i>)	1	9.76	9.76	9.38	$<.01$
D x S	1	1.01	1.01	<1.00	$>.05$
Square uniqueness	6	10.78	1.80	1.73	$>.05$
Error	36	37.65	1.04		
Total	63	145.00			

*Tested against individuals within sex and sequence; all other sources of variation tested against error.

Starting position	Means (degrees)†		Over-all mean starting position
	Position of Dangerous Edge		
	Right	Left	
30° Right (CW)	+1.19	+1.59	+1.39
30° Left (CCW)	+0.16	+1.06	+0.61
Over-all mean danger	+0.67	+1.33	

†A plus sign indicates position of the apparent vertical clockwise of the plumb line (as viewed by *S*).

compared with danger on the *S*'s right ($+1.67^\circ$). In addition, the effect of starting position is the same as that obtained in a number of previous experiments, namely, the apparent vertical is relatively close to the position in which the rod was set initially for the trial: the physical location of the apparent vertical is relatively to the right ($+1.39^\circ$) when the starting position is 30° right, compared with the starting position at 30° left ($+0.61^\circ$). Both of these effects, that of danger and of starting position, are significant.

The effect of simultaneous operation of both factors can be seen in the means for the four test conditions. The results are in keeping with the hypothesis that the two factors operate co-actively: the greatest shift of the apparent vertical to the right ($+1.59^\circ$) occurs when each of the two factors

singly make for a shift to the right (danger left, starting position 30° right), and the place of the apparent vertical farthest to the left ($+0.16$) occurs when each of the two factors singly make for a shift to the left (danger right, starting position 30° left); the two remaining test conditions, where the single factors have potencies in opposite directions, fall between these extremes.

F. DISCUSSION

The experiments reported in this study yield three main results on the effect of danger on space localization: (a) Under conditions of danger, asymmetrially induced by a precipice to one side, the physical position of the apparent median plane, relatively, shifts to the side opposite the location of danger; (b) analogously, under these danger conditions, the physical position of apparent vertical is rotated relatively opposite the location of danger;³ (c) there is evidence that this emotional factor of danger operates co-actively with non-emotional factors that are known to affect space localization. In the experiment on the median plane (Experiment 2), the non-emotional factor is asymmetrical placement of objects in relation to the observer, i.e., a factor which induces a displacement of the physical location of the apparent median plane toward the center of the figure; in the experiment on verticality (Experiment 3), this non-emotional factor is starting position, i.e., a factor which induces a shift of the apparent vertical in the direction of initial tilt of the rod.

In order to understand these effects of danger on localization in space within the framework of sensory-tonic theory, we make the assumption of equivalence of an emotional stimulus and an extraneous stimulus of a non-emotional kind. In the light of this conceptualization, first let us consider the effects of extraneous stimulation on space localization and the mechanism evoked by sensory-tonic field theory for interpreting these effects; we will then state the assumption necessary for generalizing our interpretation so that this mechanism of extraneous stimulation applies to the effects of danger.

We have, on the basis of evidence coming from pathology, assumed that asymmetrical stimulation coming from a source other than the object attended (e.g., light, sound, head torsion, etc.), arouses a primary tendency of the body to fall or turn in the direction of stimulation, and further, that this primary sensory-tonic innervation is counteracted by forces to the opposite side.

This information has been utilized in developing a diagram which pic-

³The first two generalizations imply the assumption that systematic trends in positions of apparent vertical and apparent median plane to left or right occur depending on the individual; we have found such systematic trends previously, and ascribed them to "individual equilibria" (8).

torially symbolizes, more specifically, the mechanisms involved in the effect of extraneous stimulation on apparent verticality and the median plane. Hypothetical diagrams of this sort are presented in Figure 1 for both the perception of verticality and the straight-ahead. Specific assumptions underlying the diagram are: (a) When maintaining a position of head tilt or head torsion to one side, counterbalancing forces ensue to the opposite side. For instance, to maintain head tilt to the left, forces to counteract gravitational pull come into play; in other instances such forces may ensue from intra-organismic righting tendencies, etc. (b) For a short-hand schematic characterization of different organismic states we use the term "equilibril axis," which simply represents the distribution of forces (innervation pattern). Thus, for example, when the body (head) is tilted to one side, falling is counteracted by muscular-tonic strains in the opposite direction; the changed distribution of forces is indicated by a shift of the equilibril axis in the direction of the counterbalancing forces; because of the changes in the organismic state (from o_x to o_y), the physical position of a line, to be perceived as vertical, has to be shifted (from s_x to s_y); the empirical evidence is that the line has to be tilted in the direction of the counteractive force in order to be seen as vertical.

To account for analogous shifts occurring with respect to the median plane, we note that if the head is turned, e.g., to the left, there is induced a turning tendency of the rest of the body in the same direction. Evidence for such a tonic change in the trunk induced by head turning may be found in neurological studies, such as those by Goldstein (1) and Schilder (3); for instance, it has been observed that in cerebellar patients where counterbalancing is impaired, with turning of the head there is induced a turning of the trunk around the longitudinal axis (1). Similarly, when the eyes are turned there is a tendency to turn the head. In normals this turning tendency, when not acted out, is balanced by counteractive forces in the opposite direction; under these circumstances, analogous to our schema for verticality, the medial equilibril axis is represented as shifting to the side of the counteractive force. According to our interpretation, the observed positional shift of the apparent median plane opposite to head turning or eye turning is a reflection of this change in the sensory-tonic state of the organism.

So much for the mechanism for the effects of extraneous stimulation. In order to translate this conceptualization so that it applies to the effects of danger coming from one side, one additional assumption has to be made. We assume that the presence of a precipice to the left or right generally arouses a fear of falling, either to left or to the right, which is considered equivalent

in its organismic effect to the somato-tonic innervation evoked by asymmetrically applied non-emotional stimuli. Analogous to the righting regulation, with fear of falling there is aroused a counteractive force to the opposite side. Under the assumption that there are balancing or counteractive innervations stipulated in the case of emotional and non-emotional stimuli, the schema (Figure 1) may be applied also for the effects of danger. It is consistent with this schematization that, in fact, the apparent median plane

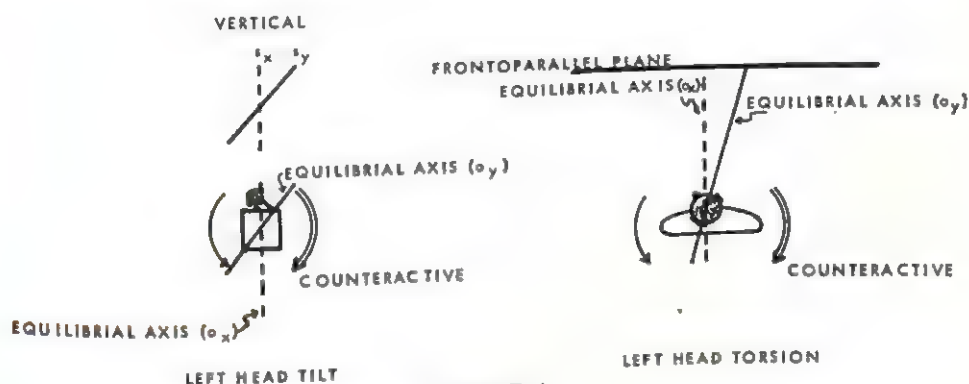


FIGURE 1

HYPOTHETICAL SCHEMA

Effect of extraneous stimulation: Position of apparent vertical under left head tilt; position of apparent median plane under left head torsion.

and apparent vertical shift to a physical position opposite the side to which the danger is located.⁴

A few comments should be added concerning the evidence on simultaneous operation of non-emotional and emotional factors introduced in Experiments 2 and 3. The non-emotional stimulation in these experiments is what we have previously called object stimulation, i.e., stimulation coming from objects, such as asymmetrical placement, evoke tendencies toward symmetrization, which, in our view, consists of shifts of sensory-tonic equilibria in the organism. These effects in our experiments are coupled with those coming from emotional stimuli, viz., danger. As discussed above, such emotional

⁴For those who think of emotions in neurophysiological terms, namely as autonomic responses, it should be noted that there is now a considerable body of anatomical and physiological evidence for the interconnection of autonomic and somatic activity. We may here refer to a statement by Kennard: "Since a large part of our knowledge of what constitutes an emotion is through observation of autonomic changes which follow it, there can be no doubt that these affective reactions have their organic counterpart in the autonomic system, and it is then reasonable to suppose that emotions may react similarly with the somatic motor system" (2). See also Teuber's review (4) and particularly his reference to Kaada's work on somatomotor and visceral effects of electrical stimulation.

stimuli are assumed to induce shifts of sensory-tonic equilibria in the organism: thus, in the last analysis, both emotional and non-emotional stimuli induce organismic changes of an identical kind. This identity in nature of organismic change appears to us as the basis for "functional equivalence" as well as "coactive operation" of emotional and non-emotional factors.

G. SUMMARY

Three experiments were carried out to study the effect of danger on space localization. Danger was injected into the experiments by placing the subject at the left or right edge of a platform elevated 31 in. from the floor. Three results were obtained: (a) Under conditions of danger, asymmetrically induced by the precipice to one side, the physical position of the apparent median plane, relatively, shifts to the side opposite the location of danger; (b) Analogously, under these danger conditions, the physical position of apparent vertical is rotated in a direction opposite the location of danger; (c) There is evidence that this emotional factor of danger operates co-actively with non-emotional factors that are known to affect space localization.

These effects were interpreted within the framework of the sensory-tonic field theory of perception.

REFERENCES

1. GOLDSTEIN, K. Über induzierte Veränderungen des "Tonus." *Schweiz. Arch. Neurol. u. Psychiat.*, 1926, 17, 203-228.
2. KENNARD, M. A. Autonomic interrelations with the somatic nervous system. *Psychosomat. Med.*, 1947, 9, 29-36.
3. SCHILDER, P. *Brain and Personality*. New York: International Univ. Press, 1951.
4. TEUBER, HANS-LUKAS. Physiological psychology. In *Annual Review of Psychology*, Vol. 6. Stanford, California: Annual Reviews, Inc., 1955.
5. WAPNER, S., WERNER, H., BRUELL, J. H., & GOLDSTEIN, A. G. Experiments on sensory-tonic field theory of perception: VII. Effect of asymmetrical extent and starting positions of figures on the visual apparent median plane. *J. Exp. Psychol.*, 1953, 46, 300-307.
6. WERNER, H., & WAPNER, S. Changes in psychological distance under conditions of danger. *J. Personal.*, 1955, 24, 153-167.
7. ———. Experiments on sensory-tonic field theory of perception: IV. Effect of initial position of a rod on apparent verticality. *J. Exp. Psychol.*, 1952, 43, 68-74.
8. ———. Toward a general theory of perception. *Psychol. Rev.*, 1952, 59, 324-338.
9. WERNER, H., WAPNER, S., & BRUELL, J. H. Experiments on sensory-tonic field theory of perception: VI. Effect of position of head, eyes, and of object on position of the apparent median plane. *J. Exp. Psychol.*, 1953, 46, 293-299.

Department of Psychology

Clark University

Worcester 10, Massachusetts

THE INTERACTION CHRONOGRAPH AS AN INSTRUMENT FOR OBJECTIVE MEASUREMENT OF INTERACTION PATTERNS DURING INTERVIEWS*¹

Massachusetts General Hospital and Harvard Medical School

JOSEPH D. MATARAZZO, GEORGE SASLOW, AND RUTH G. MATARAZZO

A. INTRODUCTION

After a clear and critical presentation of various approaches to the study of personality, MacKinnon concluded:

No longer can there be any doubt that there is both specificity and generality of behavior. Both personal consistency and inconsistency must be recognized.

Proponents of a radical anti-trait theory who deny any generality of behavior or any consistency of personality commit the "situation error," the error of assuming that all behavior is determined solely by the specific situation, physical and social, in which the individual finds himself. On the other hand, supporters of an extreme trait theory who deny any specificity of behavior or any inconsistency of personality commit the "organism error," the error of thinking of behavioral traits as fixed attributes of an *organism* as stable and as unchanging as a finger print or as a birth mark (Murphy and Murphy, 1931). A field theory of personality which sees behavior and personality as functions of a total field of which they are subparts is the form of theory which today seems best suited for the conceptual representation of personality. Such a theory is slowly emerging from the work of many investigators . . . (17, p. 43).

The problem of personality assessment has not been the exclusive concern of psychologists. For a number of years other behavior scientists, notably Moreno (18), Chapple (5), H. S. Sullivan (28), Bales (1), to mention but a few, have concerned themselves with this problem. Although the terms used ("sociometry," "interaction theory," "interpersonal relations," "group interaction and social communication") differ from those used by psychologists, the problem area is the same. An excellent review of some of these approaches will be found in Heyns and Lippitt (14), Lindsey and Borgatta (15), and Maccoby and Maccoby (16). The present article surveys Chap-

*Received in the Editorial Office on December 31, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹This study was supported by a research grant (M-1107) from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service.

ple's contributions to the subject since these have been published in a wide variety of journals not usually consulted by psychologists.

B. THE CHAPPLE INTERACTION THEORY OF PERSONALITY

For a number of years Chapple, an anthropologist, has been attempting to develop a technique for the objective measurement and description of "personality" based on an analysis in time units of the interaction between an interviewer and an interviewee (2, 3, 4, 5, 6, 7, 8). Chapple has stated: "... if we want to predict how people *will* act, the way to do it is to watch how they do act and not to infer their behavior from what they say without any means of observational check" (7, p. 199).

This approach probably reflects Chapple's (and others') dissatisfaction with attempts to evaluate "personality" characteristics in terms of second and third order inferences from the basic or raw units of observation; such as assessing oral-dependent needs from a patient's free associations, or inferring "passivity" from a patient's use of shading nuances on the Rorschach. Chapple states further that the "content" of the interview, what the patient says, is a much more difficult variable to quantify than is *how he acts* while saying it. The latter is an observable event which can be recorded by an observer and the question of "what the patient *meant* by what he said," i.e., the problem of low inter-judge reliability in interpretation of levels of explanation of the content, is avoided.²

Goldman-Eisler, an English psychologist who has used the Chapple technique in a series of recent studies (9, 10, 11, 12), has this to say regarding the problem under discussion:

There are two kinds of data with which we have to deal in conversations: those which may be summarized as (a) *what* the speaker says, and (b) *how* he says it. The first is the content, data of a highly qualitative kind which, in spite of many brave attempts at classification, tend to prove unwieldy and unruly in the process of being forced into categories. The second is concerned with the forms in which this content is presented: the expressive movements accompanying it, gestures, facial expressions, and mannerisms, the speed of talking, the pauses, silences, interruptions, the rate of talking and so on.

Experienced interviewers, for instance, pay great attention to the clues provided by these formal aspects, i.e., to the "how"—as distinct from the "what"—aspects of the things said. A number of these aspects are

²The research of several investigators, notably Carl Rogers' group, published since Chapple first made this observation indicates that the "content" of interviews can be approached in a manner which, while difficult to execute, is of the same order of scientific merit as is the study of other phenomena (21).

functions of time, which is an objectively measurable quantity. Therefore whoever wants to explore this confused field with the purpose of studying it scientifically, might in the first place profitably concentrate on just this aspect of social behavior. How long at a time a person speaks without a break or remains silent during the courses of the interview, how frequently he starts to speak and falls back in silence, how often he interrupts his partner and for how long at a stretch, how long and how frequently the persons engaged in a conversation or interview are mutually silent or else both talking at once—these can all be measured in units of time.

These data and their various relations were claimed by Chapple to be significant for the diagnosis of certain personality traits as well as of the various mental disorders (9, p. 355).

The basis of Chapple's interaction method is an analysis of the time variable during the interview. Chapple arrived at his conclusion that time was an important variable for describing human relations after considerable work in the field. He and his early collaborator, Arensberg, found that their field work as anthropologists was unduly hampered by the lack of precision and communicability of the various "subjective" variables which anthropologists (and other behavior scientists) were then using to describe human relations, in the family, tribe, interview-situation, etc. Chapple has described in everyday terms the importance of the time variable in human interactions.

When we had come to the conclusion that existing methods of appraising personality were inadequate, we decided to try a different approach. What was needed was an objective yardstick. We began, therefore, by agreeing to limit ourselves to those aspects of a person's behavior which could be directly observed and recorded. From an examination of our previous studies in evaluation of personality, we concluded that one measurable factor that seemed highly significant was time. The question then arose: what traits of personality express themselves in time (7, p. 199).

... the class of phenomena with which we are concerned comprises the (timing of) actions and interactions of individuals. It is now necessary to give a more precise definition of what this includes. As a matter of everyday observation, we see individuals coming together, and from the evidence of what we see and hear we unconsciously make certain judgments about their behavior. Such judgments are that one individual started to talk, and that the second individual to whom he was talking replied, and that both accompanied their speech with facial or bodily gestures (5, pp. 21-22).

We all know, as a matter of observation, that people have different rates (timing) of interaction. Some of our friends or acquaintances seem to talk and act very speedily as compared to ourselves; others are slow and deliberate. These characteristics of individuals are something

we intuitively recognize, and we often are at variance with the rates at which others act. For example, where there are two persons in interaction, one whose actions are quick and speech voluble, and the other, slow and given to long, well-rounded periods, we are apt to find that the speedy one keeps interrupting the slow one, jumping in when the other pauses, and so on. If the slow individual is persistent, he may finally wear the other down, and our fast individual will subside into silence broken with a few "impatient" or "bored" remarks. Or conversely, the speed at which the fast person acts may so upset the slow individual that it will throw him off his stride and he will later confess that he thought the other person "hard to talk to," "never stuck to the subject," "always interrupted" (5, pp. 31-32).

As we explained before, we do not infer and then attempt to record "feeling states" or "emotions" because we have no operations to deal with them and because we shall find that the quantitative analysis of interaction will in large measure describe such phenomena. In operations describing the timing of actions and events, we hope the reader will discover a useful and highly supple instrument. If the reader sharpens his powers of observation, he will see that in many cases people whom he does not like or cannot get along with, *say* exactly the same things that the people he does like say. So actors frequently take a short play, play it first as a tragedy and then, using the same words, play it as a comedy. Here the language is seen as unimportant, and the timing is the factor which makes the difference in its effect on the audience (5, p. 33).

From observation, we note how different people have different rates at which they originate (initiate) action. We all know cases of "bashful" people who will never speak unless spoken to. If spoken to they very frequently turn out to be very lively interactors. On the other hand, the "glad-hand artist," the "greeter," the man who speaks to everyone is a man with a very high origin (initiative) rate (5, p. 43).

The above quotations will suffice to give an introduction to Chapple's interaction theory of personality. He has taken the position that personality can be assessed without recourse to intra-psychic and other currently popular psychodynamic formulations, and further that this assessment involves merely the process of observing the *time relations* in the interaction patterns of people. Accordingly, Chapple has indicated that this method, because of its objectivity, can lead to a *science* of personality. This view, it can be seen, is consistent with MacKinnon's conclusion that the most promising approach to personality assessment will come from a "field theory" which gives sufficient weight both to "organismic" factors (the individual's behavior) and the "situational" or "field" (which involves the other interactees) variables (17, p. 43). Interestingly, Sarason has recently attempted to view the Rorschach test in a similar manner and has emphasized the examiner-subject

relationship and the effects of this on the subject's productions (23). Over the years and in work in both psychiatric settings (6) and industry (3, 4, 7), Chapple has continued to develop both his theory and the instrument of measurement, the Interaction Chronograph, required to make the necessary observations.

C. THE CHAPPLE INTERACTION CHRONOGRAPH

A history of the development of various early forms of this instrument will be found in the paper by Chapple and Donald (7); while a description and photograph of the more elaborate present instrument will be found in a recent publication by Chapple (3).

Originally, Chapple used a simple stop watch as his instrument. With this he recorded the duration of contacts between pairs of individuals (two individuals talking together, or one handing the other a tool, or some similar interaction) and recorded the results on paper ruled off in divisions for each minute that passed. He measured the length of each contact in minutes and also the length of time between contacts, and from those measurements he and his collaborator characterized interpersonal relationships in terms of *frequency of contact*, *average duration of contact*, *total time spent together*, etc. They also recorded which of the two persons *initiated* the contact. These observations were made of interactions taking place in factories (both worker and executive interactions), in a department store, and in several political organizations.

In order to get more accurate observations and at the same time to reduce the burden of the observer, Chapple and those who have worked with him next developed a simple recording device. The latter consisted of a moving tape, driven at a uniform speed, upon which lines were drawn continuously, when the keys were pressed down. When an individual started to act, e.g., in initiating a contact, the observer would depress the key assigned to that person. When he stopped acting, his key would be released and the line would end. The blank space between the lines provided a measure of his inactivity. The same procedure was used for each person under observation.

With this innovation, Chapple could measure not only the length of time each person talked or acted (his action) and the length of time he was silent (silence variable), obtaining a continuous record of this sequence of action and silence, but he could also record variations in the *adjustment* of the participants to each other. The record indicated who interrupted, how long the interruption lasted, which one *dominated* by talking the other down. Further, if there was a silence during which one person failed to respond

to the other, the record indicated who was responsible for the silence, how long the silence lasted (i.e., who was failing to "adjust" and how long this failure of adjustment lasted) and who finally *initiated* the next action.

Chapple could do more with his instrument than record these basic elements of the pattern of give-and-take in a conversation or psychiatric interview. If one person tended to dominate the interview, taking the greater portion of the time and, when interrupted, increasing the length of each speech in order to out-talk the other person, it could be determined with accuracy how the other person reacted. Did he tend to become more silent as the other talked more and more, or, refusing to become completely dominated, did he increase his own activity and try to hold his own? Measurements could also be made of his degree of adjustment to the person who was doing most of the talking. Did he interrupt or fail to respond, and if so, how frequently and for what duration?

These data and cumulative curves drawn from them were arrived at only after laborious measurement by ruler of the lines which were drawn on the moving tape. In order to reduce the tediousness of the measurements, Chapple next developed a computer which would draw cumulative curves of the interaction patterns while these were being recorded by the observer; i.e., much as Skinner's animals have always plotted their own learning curves. When the interview ended, a series of graphs had been drawn on a wide sheet of paper and, by the use of a plexiglass scale, complete measurements of the desired variables (the durations, the slopes or rates of activity, etc.) could be obtained from these graphs. This new computer was given the name Interaction Chronograph, and was the model being used in 1946 (7).

Though used by Chapple for personnel selection and research in industry for several years (3, 7), this model was limited in several ways: it did not give measurements of all the variables thought useful for even one person, and secondly, it was not designed to yield completely independent values for the interviewer. Moreover, it seemed to Chapple, since the final step in scoring was to obtain numerical values for the observed interaction, that it would be desirable to develop a machine which, besides recording the interaction of both persons completely, converted the raw measurements directly into the desired numerical data. In this way graphs would be eliminated, scoring time would be reduced immensely, and the obtained scores would be less vulnerable to scorer-errors.

The new model of the Interaction Chronograph, a photograph of which was published in 1949 (3), has not been modified further and is the machine

currently in use.³ It consists of two parts: a small observer's box and a large recording machine with which the former is connected by a long, flexible cable, thus providing for the possibility of the two parts being housed in different rooms.⁴

The observer's box has five keys or buttons which are operated manually by the observer while the interview is in progress. Two of these buttons control the power and turn the larger, recording machine on and off. A third button is a signal marker which activates a designated electrical counter in the recording machine. The purpose of this signal counter is to mark the introduction of a change at a particular point during the interview. The two remaining keys are designated, one each, for the interviewer and interviewee. After turning on the two power keys and pushing the signal button to indicate that this is the start of a new observation, the observer has only to keep one finger on the interviewer's key and one finger from his other hand on the subject's key and thus record the subsequent interaction. Each time the designated person starts to act by talking, nodding the head, gesturing or communicating in any other form, the observer presses his key and keeps it down; releasing the key only when the action comes to an end. Thus the observer's task is to press a key when an action begins, record how long it continues, and when it terminates. Since this kind of record is kept for both persons simultaneously in the situation, a complete record of their interaction is available.

The Interaction Chronograph machine, which records the interactions, is nothing more than a very elaborate electrical stop watch. The tedious computations which had to be made by the observer when Chapple was using only a stop watch (durations of action, frequency of interruptions, duration of interruption, etc.) are now all computed electrically and simultaneously, simply as a function of the combinations which are possible from an analysis of the positions of the two keys, interviewer and subject, relative to each other.

D. THE INTERACTION CHRONOGRAPH VARIABLES

There are currently 10 counters (variables) based on the relative positions of the interviewer and interviewee keys which the machine records during interaction. Some of these may seem arbitrary since they represent algebraic sums of two variables rather than individual measures of each of

³For information regarding this machine, the manual of instructions for scoring, etc., write to Dr. E. D. Chapple, 20 Woodland Drive, Noroton, Connecticut.

⁴Chapple also has a small, portable instrument for use outside of structured interview situations.

these variables. Apparently Chapple, in developing his interaction theory, found these derived variables more useful than the first-order variables from which they were obtained. The 10 variables, in the order in which they are printed by the machine and excluding the signal counter which is the first, are from left to right:

1. *A's Units*: The designation "*A*" is used rather than "Pt.'s units" since interviewees other than patients are often recorded. This counter provides a *frequency* count of the actions of the interviewee. Each time he is active, and the observer depresses his key, the unit counter adds one unit. Thus a cumulative record is made of the number of times the individual was active during a period designated by the signal counter. In view of the fact that different people have different units of action during comparable interviews or time periods, Chapple divides each of the following variables (2-9) by the number of *A's* units and thus obtains mean values which make possible comparisons between individuals.

2. *Tempo*: The present machine does not record interviewee actions and silences separately. (These, however, can be computed readily when desired.) Rather, this counter records the *duration* of each action *plus* its following inaction in a single measure. Thus it provides an index of how often one starts to act; the duration from one action to his next action; or his "tempo" of starts.

3. *Activity*: This, too, is a *duration* measure derived from an individual's actions and silences. However, unlike Tempo which is a measure of actions *plus* silences, the activity counter adds actions and *subtracts* silences. The cumulative record indicates, therefore, for any interview sequence, how much more active a person was than he was silent. Thus, the activity counter is an index of a subject's "energy level," states Chapple.

4. *A's Adjustment*: This counter, unlike Tempo and Activity, does not operate all the time the machine is on. It records only *durations* during which the person *interrupted B* (the other person) or *failed to respond to B* following the latter's last action, both of which are failures of adjustment by *A*. The counter adds the durations of interruptions and subtracts from these the durations of failures to respond. Thus a positive "adjustment" score indicates that on the average *A* interrupted longer than he was silent. A negative score indicates the converse. One would expect a type of "manic" subject to earn a positive score while some "depressed" individuals should earn negative adjustment scores.

5. *B's Adjustment*: The other individual's adjustment (i.e., the adjustment of the interviewer, usually). For a description substitute "*B*" for "*A*" in (4) above.

6. *Initiative*: This is a *frequency* counter which records only after a period in which both *A* and *B* were silent (both keys up). It provides a measure of the relative frequency with which one person, *A*, takes the initiative as against the other person, *B*. It adds just one each time *A* takes the initiative and subtracts one each time *B* takes the initiative. A person who initiates more often than he has to be initiated to earns a plus score whereas a negative score indicates the converse.

7. *Dominance*: This is a *frequency* counter which records only during an interruption (double action, with both keys down). It provides a measure of the relative frequency with which one person, *A*, out-talks or out-acts the other person, *B*, when there has been an interruption. The Dominance counter adds one if *A* dominates *B* in a double action, or subtracts one if *B* dominates *A*. A positive score indicates that *A* was more dominant than *B* in these exchanges, and a negative score the converse.

8. *A's Synchronization*: This is a *frequency* counter which records the number of times *A* has interrupted *B* or failed to respond to *B*, i.e., the number of times *A* failed to synchronize with *B*. The earlier described counter, *A's* Adjustment, records the relative duration of these events; the present counter, their frequency.

9. *B's Synchronization*: The other individual's synchronization (i.e., the synchronization of the interviewer, usually). For a description substitute "*B*" for "*A*" in (8) above.

10. *B's Units*: This counter provides a *frequency* count of the actions of the other person (usually the interviewer). It adds one unit each time the observer depresses *B's* key during an interactional exchange.

Values for these 10 variables can be read directly off the printed machine record. Individual variables such as the specific values for actions, silences, number of interruptions, initiatives, failures to respond, etc., can be abstracted from the algebraic sums without too much difficulty once the scorer has become familiar with the variables. In addition to these already mentioned, and also other "personality" variables, Chapple has recently developed a number of derived "temperament" variables such as "pretentiousness," "emotionality," "ability to take the floor," "shyness," and some 15 or 20 others (8). These are obtained by using various arithmetic combinations, under specified interview conditions, of the 10 variables described above. Thus they represent second and third-order variables. As a method of personality assessment they represent a new and fresh approach. Considerable research will be required, however, before the "validity" or "predictive meaning" of any of these variables can be ascertained. Saslow and Matarazzo

and their associates are currently attempting to study some aspects of the validity of these variables in a large scale research analysis of the psychiatric interview being carried on at the Massachusetts General Hospital.

E. THE INFLUENCE OF INTERACTEES UPON EACH OTHER DURING INTERVIEWS

In his early study of interaction patterns during interviews, Chapple placed little restriction upon the interviewer other than that he should use a non-directive or counseling interview of the type described by Rogers (20). Chapple soon discovered, however, that every interviewer was different not only in the way he behaved but in the results that he obtained from the same subject (3, p. 8; 4, p. 23; 7, p. 203). Thus it became clear not only that different interviewers have different interaction patterns when behaving in their own characteristic manner but that, as a result of these interviewer differences, different interaction patterns were elicited from the same patient when seen by two different interviewers. This was apparently true even though one could perceive no difference in what the interviewers said as judged by stenographic transcripts and even though they were following a supposedly uniform style of interviewing. Note the similarity between this and the recent finding by Raines and Rohrer (19) that differences in impressions of the *same* patient by different psychiatrists are due to personality differences in the psychiatrists, themselves.

Chapple's evidence for this extremely subtle, but theoretically very important point, came from two reliability studies he conducted (3, pp. 8-9). The first was on 40 men who were each interviewed (independently) by two psychiatrists, each of whom used his own interview pattern; and the second series was on 12 persons interviewed under similar conditions except that the two interviewers behaved in a standardized manner (to be described below). The Pearson reliability coefficients for the Tempo and Activity scores went from .53 and .79, respectively, in the first study to .93 and .95 in the second. Thus, the effect of standardizing certain features of the interviewer's behavior which hitherto had not been dealt with explicitly was to elicit almost identical patterns (for these two variables) from individuals interacting with two different interviewers.

An equally striking demonstration of the possible interviewer-effect on the conversation pattern of the second person becomes evident if one re-analyzes some data Chapple presented in 1940 (5, p. 35). Chapple gives mean values for the action variable during conversation for Individuals *A* and *B* who met together in interaction eight times over a period of two months. The

range of *A*'s daily average duration of action during these eight conversations was 3.10 to 7.81 hundredths of a minute, while *B*'s actions ranged from 2.89 to 4.33. If one computes a rank order correlation from the data Chapple presents in his table, i.e., the eight daily values for *A*'s action as against the eight paired values for *B*, one finds a striking value for ρ of .83. Thus *A* and *B* were having a subtle but marked effect on each other's duration of utterances. This finding has considerable implication for planned behavior therapy, i.e., an interviewer might, by plan, increase his own average action with low activity patients and decrease it with patients with abnormally high activity, and maintain this "prescribed" behavior therapy over a planned number of months. (The reader will note certain similarities between this and Skinner's method of operant conditioning.) Saslow and Matarazzo and their associates are currently doing research on this phenomenon.

That experienced therapists can be flexible in their interaction patterns with different subjects and thus, theoretically at least, practice prescribed behavior therapy was demonstrated in a recent study by Saslow, Goodrich, and Stein (26). These investigators studied, by an old form of the Interaction Chronograph, the behavior of a single therapist with 12 psychiatric patients he was interviewing, and in 7 conversations with three of his colleagues. The range of interaction behaviors available to this therapist, as well as the significant differences in his behavior with the two groups, are well illustrated in this study.

Goldman-Eisler, working at Maudsley Hospital in London with an ink-writing version of Chapple's Interaction Chronograph, recently has added further evidence for the influence of interviewers' behavior on the behavior of interviewees. In her first study (9), using seven members of the Psychology Department as subjects, she demonstrated that non-experienced conversationalists each had interaction patterns (short and long silences and short and long actions) which were rather stable for any particular individual. She also found that the silence variable (this differs a little from Chapple's) was a somewhat more stable personality characteristic than was action. Her next study was designed to investigate individual differences between interviewers (three senior psychiatrists) and the effects of these differences on the interaction patterns of patient-interviewees (10). Each of her interviewers used his own pattern of interviewing, i.e., the latter was non-standardized and was one of the variables under study. She confirmed Chapple's findings in that the three psychiatrists each had his own individual interaction pattern regardless of the type of patient he was interviewing (de-

pressed versus active patients); although this pattern could be adjusted to the type of patient (within the limits of each interviewer's own pattern) somewhat. Furthermore, her results are striking in their demonstration that these three doctors influenced the interaction patterns (total activity, and ratio of short silences to long silences) of the same 10 patients in different ways. Thus depressed patients talked more with one doctor than with another, while these same doctors had opposite effects on talkative (anxious) patients. Goldman-Eisler notes, "Thus it would seem from the above that the depressed patient responds best to active stimulation, and the one who talks easily gets the best chance with the passive interview technique" (10, p. 670). Here, again, is evidence suggesting that planned "behavior therapy" may ultimately be possible.

The similarities between these various results obtained in complex, molar (interview) situations and the laboratory findings of Greenspoon (13), and more recently of Sidowski (27) are very striking. Thus Sidowski, by the simple procedure of blinking a light as a reinforcer each time his subjects produced a plural word in a situation in which they were instructed only to "start saying words," was able to demonstrate the effect of learning "without awareness" on the part of the learner. Greenspoon obtained similar results by using the verbal response "umm-hmm" by the experimenter each time the subject gave a plural response. All of these results are provocative and considerably more research will no doubt be done in this area.

Goldman-Eisler indicated that her Interaction Chronograph personality descriptions of the three psychiatrists possessed some validity since the rank order correlations between objective values so derived and opinions obtained *independently* from seven psychiatrist colleagues "were in complete conformity" (10, p. 671). Considerably more research of this type is indicated, however. Goldman-Eisler next reported a study in which she demonstrated individual differences among such language (content) variables as word rate, self-reference terms, verb-adjective ratios, etc., among the three psychiatrists, as well as within the patient group (11). Her latest study reports differences in speed of talking among individuals (normals and patients) and also the sensitivity of this variable to the interaction characteristic (rate of utterance) of the other person (12). Thus her earlier results were confirmed for another interaction variable.

This series of studies seems to confirm MacKinnon's conclusions of "specificity and generality of behavior" (17, p. 43) and helps to define the parameters of this specificity and generality.

F. THE STANDARDIZED INTERVIEW

These experimental results also help to define some of the uncontrolled variance which in the past has made the interview (i.e., interviewer's behavior) a less than adequate research tool. Experimenters have attempted to use the interviewer as an *independent* variable in their efforts to measure various interviewee characteristics (the dependent variables). However, experience has indicated, and the above results confirm, that "the interviewer" is not an objective scientific instrument and that, in fact, interviewers are themselves dependent, in the sense of unstandardized variables. It is little wonder then, for example, that the science of psychotherapy, where the effects of this procedure have been "evaluated" by various interviewers, has made as little headway as it has (24). Furthermore, to take just one more example, the consistently low correlations in the literature, between interviewers when evaluating patients in regard to a specific diagnosis and/or other personality characteristics, may be attributable to this lack of objectivity in interviewer behavior. It may be these objectively quantifiable interactional characteristics of interviewers which are the differences in interviewer "personality" that led to such low inter-psychiatrist reliability in the study by Raines and Rohrer (19).

As early as 1946 Chapple reported an attempt to make the interviewer's behavior, as well as the interview itself, a more reliable scientific instrument (7). The consequent increase in reliability for two of his dependent variables (patient's Tempo and Activity) was discussed earlier in the present paper. Chapple's standardized method involved certain "rules" for the interviewer to follow in his own interviewing behavior (interaction pattern) and in the over-all conduct of the interview itself. The latter involved dividing an interview into five standardized periods: with Periods 1, 3, and 5 as free give-and-take periods and Periods 2 (silence) and 4 (interruption) as *stress* phases of the interview. In a recent paper giving a very detailed description of this standardized interview Chapple indicates:

By varying the interviewer's behavior, the standardized interview presents a sequence of different behavioral situations designed to elicit from the interviewee his characteristic traits of personality and temperament. In order to do this, the interview is designed to create situations involving a mild degree of stress for the interviewee, comparable to those which he may normally encounter in daily life. The stress is provided by the behavior (note, not content) of the interviewer. At different stages of the interview, the interviewer changes the pattern of behavior he has been following in order to confront the interviewee with a different type of behavior than he has previously shown. This means

that we are using the interviewer as the independent variable and are prescribing measured amounts of specific types of interactional behavior to determine its effects on the interviewee. By doing this under controlled conditions, and by measuring the reactions of the interviewee to a sequence of different behaviors, the standard interview provides information more representative of the ranges of adjustment in interaction of which the interviewee is capable than would be the case if the interviewer followed the same pattern throughout (4, p. 23).

Thus Chapple's standardized interview provides a controlled sample of patients' behaviors in miniature, molar situations. It is these molar or "field" situations which MacKinnon's earlier reported comments indicated are necessary for an adequate assessment of personality (17).

TABLE 1
CHARACTERISTICS OF THE STANDARDIZED INTERVIEW

Period	Type of interviewing	Duration of period	
		Fixed duration	Variable duration
I	Free	10 minutes	
II	Stress (Silence)		12 failures to respond, or 15 minutes, whichever is shorter
III	Free	5 minutes	
IV	Stress (Interruption)		12 interruptions, or 15 minutes, whichever is shorter
V	Free	5 minutes	
Total		20 minutes	plus a maximum of 30 more minutes

By establishing "rules" for the interviewer's behavior during the conduct of these five periods, Chapple has developed an "instrument" (the interview) which will make more comparable the results obtained by one interviewer with the same patient at different times; as well as the results obtained by different interviewers working in the same or other settings. In Tables 1 and 2 will be found a summary from various sources (3, 4, 7) of the standardized interview and the interviewer's standardized behavior.

It can be seen in Table 1 that three of the (adjustment) periods have fixed durations, whereas the durations of the two stress periods are variable. The interviewer starts off the interview with a five-second comment of the variety, "I wonder if you would mind telling me just how it was you happened to come (or were sent) here to see me." During the remainder of

Period 1 he uses a non-directive type of behavior. After 10 minutes, and to introduce all other periods, the interviewer signals the observer (this can be accomplished by pushing a button which flashes a light for the observer who is watching through a one-way screen) indicating that his next utterance will begin Period 2. During Period 2 (which will require some practice by interviewers) the interviewer fails to respond to the patient's last comment 12 consecutive times; this is most difficult for the interviewer on those rare occasions when subjects ask direct questions.

TABLE 2
STANDARDIZED INTERVIEWER'S BEHAVIOR: RULES FOR INTERVIEWER

Periods 1 to 5 (all periods)

- a. Interviewer introduces each period by a 5-second utterance (following his signal to the observer).
- b. All interviewing must be *non-directive*. No direct questioning, no probing or depth interviewing. Interviewer can reflect, ask for clarification, ask for more information, introduce a new topic area, etc. In general, interviewer's comments should be non-challenging and open-ended and related to the patient's past comments or to some new, general topic.
- c. All interactions must be verbal only, or verbal and gestural at the same time; i.e., interviewer cannot use head nods and other gestures alone. This rule simplifies the observer's task.
- d. All of interviewer's utterances must be of approximately 5 seconds duration.
- e. After patient finishes a comment or other interaction, interviewer must respond in less than 1 second, except as otherwise noted in Period 2.
- f. Each time patient interrupts interviewer, the latter must continue to talk for 2 more seconds. This rule insures more explicit definition of a patient's ascendance-submission pattern than would be possible if interviewer "submitted" immediately.

Periods 1, 3, and 5

- a. Interviewer must never interrupt patient.
- b. If after interviewer makes a comment patient does not respond, interviewer must wait 15 seconds and then speak again for 5 seconds.

Period 2 only

- a. Interviewer must "fail to respond" to last interaction of patient a total of 12 times (or for 15 minutes, whichever is shorter).
- b. After interviewer has been silent for 15 seconds (and patient has not taken initiative) interviewer makes another 5-second comment.

Period 4 only

- a. Each time patient acts, interviewer must interrupt patient for 5 seconds for a total of 12 times.
- b. Interviewer's interruption should begin about 3 seconds after patient has begun his interaction.
- c. After having interrupted patient, if the patient continues through the interruption (does not submit), interviewer will not interrupt again until patient has finished his utterance, i.e., interviewer will interrupt patient only once during each utterance of the latter if patient does not "yield."
- d. The period is ended after 12 interruptions or 15 minutes of attempting to obtain these.

Many people will speak again (take the initiative) under these "silence" conditions. If, during a failure to respond, the subject does not start again, the interviewer waits 15 seconds (this is gauged from a big sweep hand electric clock on the wall directly behind the patient) before restimulating with another five-second utterance. During this "silence" period a person's characteristic "initiative" level can be objectively assessed. Since people show wide individual differences during this period it may not always be possible for the interviewer to fail to respond 12 times. Especially is this so with individuals who have long durations (over two minutes) to each of their utterances. Therefore a limit of 15 minutes is placed on the duration of this period. Period 2 is followed by a 5-minute period of adjustment similar to Period 1. This is followed by a second stress period in which the interviewer interrupts the patient in a standardized manner for a total of 12 times or 15 minutes, whichever is the shorter. The subject's characteristic "dominance-submission" interaction pattern can be objectively assessed during this phase of the interview. Following this interruption period is a final 5-minute adjustment period. From Table 1 it will be apparent that the durations of the interview conceivably can range from slightly over the fixed duration of 20 minutes to a maximum of 50 minutes.

In regard to Table 2 it should be noted that in his 1953 paper Chapple indicates (4, p. 25) that sustaining monosyllables such as "uh-huh," "Well," etc., may be used occasionally by the interviewer. However, the experimental results presented earlier (especially the rho of .83 between the average lengths of *A*'s and *B*'s actions) would indicate that a fixed duration for each utterance (approximately 5 seconds for example) is desirable. With a stop watch and relatively little practice an interviewer should have no trouble training himself to produce 5-second utterances. For example the patient-question, "Shall I continue?" can be answered with, "Yes, in view of what you said earlier about your parents I believe it would be well if you did continue that," instead of with merely "Yes."

This attempt by Chapple to somewhat standardize interview behavior may appear artificial and restrictive. However, he apparently is following a scientific tradition which holds that, if objectivity is a requirement of a scientific instrument, some standardization is mandatory.

G. THE RÔLE OF THE OBSERVER

It will be apparent from the discussion so far that the observer is an important person in the conduct of research with the Interaction Chronograph, since he is, in effect, part of the instrument. All data are recorded and

therefore contributed by him. Therefore he must be a faithful transmitter of the interviewer-interviewee behavior. For this reason it is important that he (as well as the interviewers) go through a practice series before beginning a research enterprise utilizing this instrument. Experience indicates that a practice series of 25 to 50 interviews is probably adequate.

In general, the observer will have little difficulty deciding what is scorable action on the part of the subject or interviewer. The latter is required to use verbal communication primarily, with gestural communication always clearly secondary, thereby making him "easy to record." Chapple deals explicitly with the problem of what to score as action (4, p. 29) although he treated it in some detail in his earlier monograph (5, pp. 24-25). In the interview situation experience indicates that approximately 90 per cent of the interviewee action is also verbal, with the remaining (and more difficult to score) interviewee communication being made up of head nods, gesturing, communicative smiling or frowning, etc. A further problem occurs for the observer when he has to decide when an action has terminated. Even if one follows Chapple's general rule that all facial (and other) muscle activity must have come to an end for action to be scored as ceased, there is an occasional subject who speaks in halting phrases during a single action. In these cases, the present writers have used the verbal context as an aid for making a decision. Thus a subject who says, "While talking to my boss that time I . . ." (and then hesitates briefly before continuing, during which time all facial activity has ceased as far as the observer can see) is scored as *continuing* action or communication during this period. Differentiating between these pauses during on-going action and the "blocking" characteristic of psychotic patients, which is *not* scored as action by Chapple, is not too difficult a process after some experience.

Specific research on the question of observer reliability (i.e., reliability of the instrument, *per se*, rather than reliability of the subjects studied with it), while difficult to implement, should nevertheless be conducted. This can be done most convincingly by having two observers, using two instruments, record independently the same series of interviews, or it can be done in several other ways. Both Chapple and Goldman-Eisler have indicated their recognition of this problem. Thus Chapple notes:

Since the criteria for observation of interaction are objective and involve moving the key only when there is a visible contraction or relaxation of facial muscles, there is little chance for difference in normally skilled observers. A small series of half a dozen records done by two observers on tape recorders showed no difference in the means obtained or the number of units recorded. The present machine also

has a threshold built in so that durations of under .25 seconds do not record, and this threshold seems to be at the level of individual differences in reaction time (3, p. 9).

Referring to the same problem Goldman-Eisler wrote:

Chapple's method of interview analysis consists fundamentally in recording on a chronograph tape the durations of actions (speech and gesture) and silences of both persons involved. Apart from talking, those kinds of gestures were interpreted as actions which were of a response character or part of the social interaction process, such as nods, certain kinds of smile, or shaking of the head, and laughter. Although this is a matter of judgment, preliminary recording by two observers in the present investigation showed no essential differences in the interpretation, once the criteria were agreed upon (9, p. 355).

H. RELIABILITY AND VALIDITY OF THE INTERACTION CHRONOGRAPH METHOD

As is the case with all new research instruments the problems of reliability and validity are extremely complex. As indicated earlier, Chapple reports reliability values of .95 and .93 for interviewee Tempo and Activity, respectively, in a study of 12 persons interviewed once each by two different interviewers using the standardized interview (3, pp. 8-9). Goldman-Eisler's data on *stability of interviewee* behavior (9, 10), while not specifically concerned with the question of "reliability," can nevertheless be accepted as indirect evidence for this point. Since her *interviewers* were not using a standardized interview but were being studied as dependent variables themselves, their uniformity across different patient conditions can be similarly interpreted as indirect evidence for the reliability of the method. Finally, Saslow, Matarazzo, and Guze in an extensive recently completed study (25) of the reliability of the Interaction Chronograph, obtained test-retest reliabilities (using two different interviewers) ranging from .726 to .956 for the 10 variables defined earlier in the present review (Table 2).

The validity question is a thorny one because it involves the question, "Validity for what?" Goldman-Eisler's treatment of the validity problem was discussed above. Essentially she found "complete conformity" between her objective, chronograph (personality) descriptions of her three psychiatricists and the independent ratings (opinions) on these variables by seven colleagues (10, p. 669). In an early paper by Chapple and Lindemann (6) these investigators, using an old version of the instrument and plotting individual graphs for each patient, were able to differentiate manic patients, depressed patients, hysterics, and other psychiatric patient groups. However,

the present method of representing the results is far superior to this earlier version and more research in this area is indicated. In a more recent study (3, pp. 9-10) Chapple found values for his "rigidity" and "flexibility" variables which significantly differentiated a group of 168 psychiatric patients from a group of 258 controls ("normals").

By far the most data bearing on the problem of validity have been gathered in studies of "applied anthropology" or "applied psychology." Chapple has been applying his Interaction Chronograph method to the problem of employee selection and placement. On the basis of work completed, Chapple reports (3) that some of his indices discriminate between (a) effective and ineffective sales personnel; (b) individuals in different job groups in a company; for example, Chapple reports the degree of "initiative" needed varies in different types of selling, as does also "activity," since much higher activity is needed in a rapid-fire operation like hosiery selling than in the selling of a hat, a pair of shoes, or large household furniture (3, pp. 12-13); (c) workers and foremen, as well as workers and executives; (d) and other groups.

However, these validity studies, while adequate for reports in the field of applied psychology for which they were intended, need to be elaborated and communicated with presentation of considerably more of the results if this method is to be used for exploration and research in the field of basic personality structure. Chapple is apparently aware of this, since he devotes a section of one of his papers to a specific discussion of "The validation problem in industry" and contrasts this with the more traditional treatment of the problem of validity (3, pp. 10-11).

I. SUMMARY

Various instruments of personality assessment have been employed in the past. Research to date has shown that one of these instruments, the interview, while potentially very useful is limited because of the problem of low reliability between independent interviewers. Part of this difficulty seems to be due to the fact that *the interview*, i.e., the interviewer's behavior, has been considered by many investigators to be the *independent* variable, with elicited interviewee behaviors as the dependent variables. However, published research would indicate that *the interviewer* is not an objective scientific instrument and that, in fact, interviewers are themselves dependent, in the sense of unstandardized, variables. The present paper concerns itself with a review of the history and development of a promising new research instrument, the Chapple Interaction Chronograph, and the

attempt made by Chapple to standardize, and therefore make objective, the interview itself as a research instrument. Published research in which this method has been employed also was reviewed.

REFERENCES

1. BALES, R. F. *Interaction Process Analysis*. Cambridge, Mass.: Addison-Wesley, 1951.
2. CHAPPLE, E. D. Quantitative analysis of the interaction of individuals. *Proc. Nat. Acad. Sci.*, 1939, **25**, 58-67.
3. ———. The Interaction Chronograph; its evolution and present application. *Personnel*, 1949, **25**, 295-307.
4. ———. The standard experimental (stress) interview as used in interaction chronograph investigations. *Human Organiz.*, 1953, **12**, 23-32.
5. CHAPPLE, E. D., with ARENSBERG, C. M. Measuring human relations: An introduction to the study of the interaction of individuals. *Genet. Psychol. Monog.*, 1940, **22**, 3-147.
6. CHAPPLE, E. D., & LINDEMANN, E. Clinical implications of measurements of interaction rates in psychiatric interviews. *Appl. Anthropol.*, 1942, **1**, 1-11.
7. CHAPPLE, E. D., & DONALD, G., JR. A method for evaluating supervisory personnel. *Harvard Bus. Rev.*, 1946, **24**, 197-214.
8. CHAPPLE, E. D., CHAPPLE, M. F., & REPP, J. A. Behavioral definitions of personality and temperament characteristics. *Human Organiz.*, 1954, **13**, 34-39.
9. GOLDMAN-EISLER, F. The measurement of time sequences in conversational behavior. *Brit. J. Psychol.*, 1951, **42**, 355-362.
10. ———. Individual differences between interviewers and their effect on interviewees' conversational behavior. *J. Ment. Sci.*, 1952, **98**, 660-670.
11. ———. A study of individual differences and of interaction in the behavior of some aspects of language in interviews. *J. Ment. Sci.*, 1954, **100**, 177-197.
12. ———. On the variability of the speed of talking and its relation to the length of utterances in conversations. *Brit. J. Psychol.*, 1954, **45**, 94-107.
13. GREENSPOON, J. The effect of verbal and non-verbal stimuli on the frequency of members of two verbal response classes. Unpublished doctor's dissertation, Univ. Indiana, 1950.
14. HEYNS, R. W., & LIPPITT, R. Systematic observational techniques. In G. Lindzey (Ed.), *Handbook of Social Psychology*. Cambridge, Mass.: Addison-Wesley, 1954. (Pp. 370-404.)
15. LINDZEY, G., & BORCATT, E. F. Sociometric measurement. In G. Lindzey (Ed.), *Handbook of Social Psychology*. Cambridge, Mass.: Addison-Wesley, 1954. (Pp. 449-487.)
16. MACCOBY, E. E., & MACCOBY, N. The interview: a tool of social science. In G. Lindzey (Ed.), *Handbook of Social Psychology*. Cambridge, Mass.: Addison-Wesley, 1954. (Pp. 449-487.)
17. MACKINNON, D. W. The structure of personality. In J. McV. Hunt (Ed.), *Personality and the Behavior Disorders*. New York: Ronald Press, 1944. (Vol. 1, pp. 3-48.)
18. MORENO, J. L. *Who Shall Survive?* Washington, D. C.: Nerv. & Ment. Dis. Monog., 1934, No. 58.
19. RAINES, G. N., & ROHRER, J. H. The operational matrix of psychiatric practice: I. Consistency and variability in interview impressions of different psychiatrists. *Amer. J. Psychiat.*, 1955, **11**, 721-733.

20. ROGERS, C. R. *Counseling and Psychotherapy*. Boston: Houghton Mifflin, 1942.
21. ROGERS, C. R., & DYMOND, R. F. *Psychotherapy and Personality Change*. Chicago: Univ. Chicago Press, 1954.
22. RORSCHACH, H. *Psychodiagnostics*. New York: Grune & Stratton, 1942.
23. SARASON, S. B. *The Clinical Interaction: With Special Reference to the Rorschach*. New York: Harper, 1954.
24. SASLOW, G. *Psychotherapy*. In C. P. Stone (Ed.), *Annual Review of Psychology*. Stanford, Calif: Annual Reviews, Inc., 1954. (Vol. 5, pp. 311-336.)
25. SASLOW, G., MATARAZZO, J. D., & GUZE, S. B. The stability of interaction chronograph patterns in psychiatric interviews. *J. Consult. Psychol.*, 1955, **19**, 417-430.
26. SASLOW, G., GOODRICH, D. W., & STEIN, M. Study of therapist behavior in diagnostic interviews by means of the interaction chronograph. *J. Clin. Psychol.*, 1956, (in press).
27. SIDOWSKI, J. B. Influence of awareness of reinforcement on verbal conditioning. *J. Exp. Psychol.*, 1954, **48**, 355-360.
28. SULLIVAN, H. S. *The Interpersonal Theory of Psychiatry*. New York: Norton, 1953.

Massachusetts General Hospital
Boston 14, Massachusetts

THE COLLECTION AND USE OF RETROSPECTIVE DATA*

Department of Psychology, Michigan State University

PAUL BAKAN

A. THE PROBLEM

The definition of the stimulus is a source of confusion in psychological discussions using stimulus-response terminology. It would be convenient to assume that the stimulus is whatever the experimenter defines it to be, but this assumption would oversimplify matters. Actually, the stimulus is dually defined or constituted, by both the experimenter and the subject. Things in the environment are attended to, perceived, interpreted, evaluated, or perhaps even ignored by the subject. As a result of such activities the subject defines what is for him the stimulus and this does not necessarily correspond to the stimulus as defined by the experimenter. Yet it is the stimulus as it is constituted by the subject that is the most important determinant of his responses.

The rôle of the subject in constituting the stimulus was stressed by Dewey (3) in his paper on the reflex arc concept. Koffka (4), in his distinction between the geographical and the behavioral environment, also recognized the rôle of the subject in constituting the stimulus. More recently, several authors (1, 5) have considered some of the implications of the dual nature of the stimulus.

The distinction between the two conceptions of the stimulus, that of the subject, and that of the experimenter, has important methodological implications. If the subject's conception of the stimulus situation is an important determinant of his responses, then it is important to find out something about it. With human subjects some form of subjective or phenomenological report can be very useful in obtaining this information. It is interesting to note that the Gestaltists who made the distinction between the geographical or physical environment and the behavioral or subjective environment also advocated the use of phenomenological data in the study of the behavioral environment. The need for systematic procedures for collecting such data has been recognized in a recent paper by D. Bakan, dealing with the problem of introspection (2).

*Received in the Editorial Office on January 16, 1956, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

The purposes of this paper are to describe and exemplify a technique for the collection of retrospective data in an objective fashion, and to consider some of the possible applications of such data. The method will be demonstrated in terms of a nonsense syllable learning situation. A detailed account of this experiment will be reported elsewhere, since the primary purpose of this paper is the presentation and application of the method rather than the report of a particular experiment. Aside from the specific examples of application, the method is described in a general way and is applicable to any experimental situation.

B. THE EXPERIMENTAL SITUATION

One hundred subjects were asked to memorize a list of 20 nonsense syllables. The total list was exposed on a blackboard in four columns of five syllables for five minutes. This was followed by a two-minute recall period, during which subjects wrote down all the syllables they could remember. The number of syllables correctly recalled by each subject was his score for this task. After learning the syllables, the subjects responded anonymously to a post-experimental questionnaire consisting of 75 items.

C. COLLECTION OF RETROSPECTIVE DATA

The essential feature of the proposed method is the compilation of a set of statements, each of which can be answered *yes* or *no* (or *agree* or *disagree*). The subject replies to these statements after participating in an experimental situation, e.g., learning nonsense syllables. The statements refer to subjective aspects of the experimental situation which are considered relevant by the experimenter. Among these statements were the following:

1. I tried hard to associate the syllables with meaningful words.
2. I frequently would look away from the board or close my eyes and test myself by trying to see how many I knew.
3. Some of the syllables just seemed to go together and this helped to make them stick in my mind.
4. I found it hard to concentrate on what I was trying to memorize.
5. I found this task of learning the syllables quite enjoyable.
6. I did not try as hard as I could have because I knew that the syllables were not worth retaining, since knowing them would be of no benefit to me.

In a short period of time it was possible to gather a considerable body of data on the subjective characteristics of the experimental situation. The subjective data, however, is gathered in an "objective" fashion, since subjects reply *yes* or *no* to each statement. An *uncertain* category may, of course, be added.

The most important aspect of this procedure is the selection of the list of statements which constitutes the post-experimental questionnaire. Any information obtained by this method is determined by the selection of statements. In order to minimize the omission of statements which might yield valuable data, several sources of statements should be considered. The purposes of the experimenter will determine which statements are actually used. In the nonsense syllable experiment the major interests were the methods used to learn the syllables and the motivational variables. Consequently, the statements chosen reflect these interests.

D. SOURCES OF STATEMENTS

Three sources of possible statements for a post-experimental questionnaire will be considered. These sources are (a) the experimenter's analysis of the situation, (b) retrospections of subjects in a pilot group, and (c) literature on the subject under investigation.

1. *Experimenter's Analysis of the Situation*

The experimenter is in a good position to produce statements relevant to the variables under investigation. If he is interested in studying the relationship between motivational variables and the learning of nonsense syllables, he may ask himself this question: "If I were a subject, what motivational variables might influence my performance in this task?" In answering his question he might think of such variables as attitude toward the experimenter, level of attention required, feeling of boredom, desire to do well, feeling of competition, sense of frustration, and so on. These variables might then suggest statements for the post-experimental questionnaire. In general the experimenter's answer to the question: "What might I think or do if I were a subject?" can serve as a fruitful source of statements.

2. *Retrospection of Subjects in a Pilot Group*

A second source of relevant statements is available in the retrospections of subjects in a pilot group. Subjects, after serving in an experimental situation, may be asked to write freely about what they did, what they thought, how they felt, and what they considered important during the experiment. Such retrospections may be facilitated by the use of incomplete sentences. These written accounts can be used as a source of statements to be included in the questionnaire.

That subjects show variation in what they write about is not a disadvantage for this method. Suppose Subject A writes about his method of asso-

ciating each nonsense syllable with a real word, but writes nothing about the difficulty of the task for him. Suppose Subject *B* writes about the difficulty of the task, but doesn't mention anything about the method he used to learn the syllables. The post-experimental questionnaire can make use of the retrospections of both the work method, reported on by *A*, and the difficulty, reported on by *B*. Information on both these variables will be elicited by each subject in the subsequent experimental group. It is known that subjects may be selective in their reports because of self-instructions about what is relevant. The disadvantages of this selectivity in individual retrospection is minimized by a pooling of many people's retrospections in a series of statements. The assumption is made that a factor relevant for one individual may be relevant for others, even if the others do not spontaneously report on it. Certain people may verbalize what others may feel without verbalizing. Putting the idea in the form of a statement may stimulate explicit thinking about it.

This is akin to the procedure involved in the construction of a personality inventory along objective lines. Consider the development of a scale of anxiety. Patients judged to be anxious may report a variety of symptoms spontaneously, e.g., "I worry over money and business," "I sweat easily on cool days." Not all patients with anxiety report the same symptoms. But each symptom reported by an anxious person may conceivably also be a symptom of other anxious people. Therefore a number of such statements representing possible anxiety symptoms are assembled and the individual can respond to each statement by indicating whether the statement is true or false for him.

3. *Literature on the Subject Under Investigation*

Another source of relevant subjective variables which might be useful in the writing of statements is the literature in any given field. Returning to the example of learning nonsense syllables, suppose that a review of the literature in this area uncovers a study of the effect of ego-involvement on the learning of nonsense syllables. Suppose that it were found that experimentally induced ego-involvement affects the learning. It is possible that attitudes experimentally induced by the instructions in a given type of experimental situation may actually be variables for a particular subject even when the experiment is not designed to manipulate these variables.

E. STATISTICAL ANALYSIS OF THE DATA

One of the advantages of the method is that it produces data in a form amenable to statistical manipulation. Some of the statistical methods applicable

are the *t*-test, analysis of variance, chi square, biserial correlation, phi coefficient, tetrachoric correlation, factor analysis, and non-parametric tests. The following examples will suffice to indicate the applicability of statistical analysis.

1. Do *S*'s who respond *yes* to statement "This test was very hard," perform differently on learning nonsense syllable list than those who say, *no*?

a. Apply *t*-test between mean number of nonsense syllables learned for *yes* group and mean number of syllables learned for *no* group.

b. Compute point biserial correlation between *yes-no* variable and number of nonsense syllables learned.

These techniques are applicable only if some *S*'s respond *yes* and some respond *no*. If there is unanimity or near-unanimity of response to the item then the problem doesn't exist. The *E* then knows that difficulty (as perceived by *S*) does or does not characterize the task.

2. Do differences in stimulus presentation have an effect on reported task difficulty?

a. If there are two stimulus presentation conditions, count number of *yes* and *no* responses to the statement "The task is difficult" for each stimulus presentation condition. Set up 2 x 2 contingency table, and test for independence between *yes-no* variable and stimulus presentation variable by means of chi square with one degree of freedom.

b. If there are more than two conditions of stimulus presentation, a larger contingency table may be used and the chi square test of independence may be made with appropriate degrees of freedom.

F. SOME USES OF THE DATA COLLECTED

1. Subjective Description

The data collected by the use of this method can serve a variety of purposes. One use of the data is to obtain a description of the experimental situation from the subjects' points of view. In the nonsense syllable experiment 75 statements were ranked according to the percentage of the subjects agreeing with the statements. In Table 1 there is a listing of the five statements most frequently agreed with (replied *yes* to) and the five statements least frequently agreed with, together with the percentage agreement for each of the statements.

An item for which there was a practically even split in response was the following one: "I found it hard to concentrate on what I was trying to memorize." This statement was agreed with by 51 per cent of the subjects and disagreed with by 49 per cent of the subjects. This indicates considerable difference among the subjects in the perceived difficulty of the

learning task. Another item of this type was the following: "I just wasn't in the mood to knock myself out learning these syllables." This was agreed with by 47 per cent of the subjects and disagreed with by 53 per cent, indicating considerable difference in motivational state within the group.

TABLE 1
MOST AGREED WITH AND LEAST AGREED WITH STATEMENTS FROM POST-EXPERIMENTAL
QUESTIONNAIRE IN NONSENSE SYLLABLE EXPERIMENT

<i>Most agreed with</i>	
1. One can learn much about experiments by serving as a subject in them.	(83% agreement)
2. I had the feeling throughout the task that I wasn't using the most efficient methods to learn the syllables.	(82% agreement)
3. I hope I haven't fouled up your experiment.	(79% agreement)
4. Some of the syllables reminded me of words, or abbreviations or foreign words, and this helped make them stick in my mind.	(78% agreement)
5. I frequently would look away from the board or close my eyes and test myself by trying to see how many I knew.	(77% agreement)
<i>Least agreed with</i>	
1. I made absolutely no attempt to learn these syllables.	(1% agreement)
2. If psychologists spend their time doing experiments like this, I don't see how psychology can get anywhere.	(3% agreement)
3. I learned the first letter of each syllable down the column and this helped me to remember the syllables.	(4% agreement)
4. For some reason I just didn't like the fellow who ran this experiment.	(5% agreement)
5. I tried to make a nonsense sentence out of each group of nonsense words.	(5% agreement)

2. *Relation between Subjective and Objective Variables*

Differences in response related to a "subjective" variable may be related to differences in response on some "objective" variable. With data of the kind described in this paper it is possible to investigate relationships between subjective variables like feeling of fatigue, desire to do well, reported work methods, and an objective variable like the number of nonsense syllables correctly recalled.

For example, do subjects who report difficulty in concentrating on the syllables learn fewer syllables than those who do not report such difficulty? This question was answered in the following manner. Subjects were divided into two groups, those who replied *yes* to the item, "I found it hard to concentrate on what I was trying to memorize," and those who replied *no* to this item. The mean number of syllables correctly recalled by the *yes* group was 7.14, and the mean number correctly recalled by the *no* group was 8.82.

The significance of the difference between these means, when evaluated by a *t*-test, was established at the 1 per cent level.

Another question which can be answered by analysis of the data is whether active rehearsal during learning is related to the amount learned. Subjects again were divided into a *yes* and a *no* group, this time on the basis of their reply to the statement "I frequently would look away from the board or close my eyes and test myself by trying to see how many I knew." The mean number of syllables correctly recalled was 8.48 for the *yes* group, and 6.22 for the *no* group. This difference was evaluated by means of a *t*-test and was found to be significant at the 1 per cent level.

Consider next the statement "I did not try as hard as I could have because I knew that the syllables were not worth retaining, since knowing them would be of no benefit to me." The mean number of syllables recalled correctly was 6.74 for the *yes* group and 8.62 for the *no* group. This difference was significant at the 1 per cent level.

3. *Discovery of Significant Variables*

An important application of the method is in the discovery of significant variables in a given situation. Such variables might then be isolated and perhaps further manipulated to give fuller insights into the phenomenon under investigation. One of the variables shown to be significant in the nonsense syllable learning situation was the presence of competitive feeling. It seems likely that this variable may be important in other group experiments as well. Subjects replying *yes* to the statement "I had the feeling that I was in a competitive situation and was trying to do better than the others," recalled significantly more syllables than subjects replying *no* to this statement. As a result of this finding one may entertain the hypothesis that explicit manipulation of the competition variable would effect the learning of nonsense syllables. Another variable discovered in this study is of interest to those who conduct experiments during class time. It was found that subjects who replied *yes* to the statement "I appreciate any break in the class routine," recalled significantly more syllables than those who replied *no* to this statement.

4. *Some Other Applications*

The method can be used in the study of relationships between responses to one statement and responses to other statements. For example, one might be interested in the relationship between feeling of fatigue and feeling of boredom in a given situation. Do subjects who report feeling fatigued tend to report that they also feel bored? With data in the form of *yes* and

no responses to relevant items such a question can be answered by means of contingency statistics obtained from a two by two contingency table.

The information gathered with this method can be used in the solution of certain practical problems. An important class of problems is the teaching of skills. Assume that for some reason we wish to teach people to become more proficient in the learning of nonsense syllables. By means of the suggested method for collection of retrospective data we have found that subjects who reply *yes* to the statement "I frequently would look away from the board or close my eyes and test myself by trying to see how many I knew" learned significantly more syllables than subjects who replied *no* to this item. Perhaps, if subjects were explicitly instructed to do this it would result in better learning. Similarly, activities characteristic of poor learners could be explicitly discouraged, and this might also lead to better learning. This information can thus be used in solving the problem of what to teach people if we want them to become more proficient in the learning of nonsense syllables.

Another application is in the study of changes in subjective factors, as a function of variation in the experimental situation. One might wish to answer questions of the following sort: "Does knowledge of results have an effect on the interest of a task?" "Does the length of an experimental session have an effect on the attitude of the subject toward the experimenter?"

The method is also applicable in situations where a given response may be made for different reasons by different subjects or where different responses may be made for the same reasons. Responses to appropriately selected items might help to get at such subjective variables.

The various applications of the method considered above are by no means exhaustive. They are merely suggestive of the types of problems that can be attacked by the use of retrospective data collected systematically and in what may be considered an objective manner.

G. OTHER TYPES OF ITEMS

Though the exposition of the method presented above has been concerned with items to which a *yes* or *no* response is called for, this is not the only type of item which can profitably be used. In some cases the use of scales rather than dichotomous responses might be more advantageous. This is especially true when the subjective variable under investigation can be thought of as having a dimensional character. The experimenter might be interested in getting reports on degree of interest, degree of effort, degree of boredom, etc. Subjects might indicate amount or degree on a linear scale.

ness of the limitations of methods of subjective report. It is hoped that the advantages of the method will outweigh its limitations.

I. SUMMARY

This paper describes a method for the collection of retrospective data in an objective manner, and demonstrates the application of the method in a study of nonsense-syllable learning. A list of statements, referring to the experience of the *S* while learning the syllables, is presented to him immediately after recall of the syllables. *S* responds to each statement by checking *yes* or *no* on the basis of the applicability of the statement to his experience. Material included in the list of statements may be derived from a variety of sources, including *E*'s own analysis of the experimental situation, the unstructured introspections of *S*'s in a pilot study, and a survey of the literature. The application of statistical analysis to the data obtained with this method is discussed and exemplified with data from the nonsense-syllable study. Some of the uses of data collected are considered. These are (a) to obtain a description of an experimental situation as experienced by the *S*, (b) to study the relationships between subjective and objective variables, (c) to discover significant variables, (d) to study relationships between subjective variables, and (e) to study the effect of experimental manipulations on subjective variables.

REFERENCES

1. BAKAN, D. Learning and the scientific enterprise. *Psychol. Rev.*, 1953, **60**, 45-49.
2. ———. A reconsideration of the problem of introspection. *Psychol. Bull.*, 1954, **51**, 105-118.
3. DEWEY, J. The reflex arc concept in psychology. *Psychol. Rev.*, 1896, **3**, 357-370.
4. KOFFKA, K. Principles of Gestalt Psychology. New York: Harcourt Brace, 1935.
5. SLACK, C. W. Feedback theory and the reflex arc concept. *Psychol. Rev.*, 1955, **62**, 263-267.
6. STEPHENSON, W. The Study of Behavior. Chicago: Univ. Chicago Press, 1953.

Department of Psychology
Michigan State University
East Lansing, Michigan

THE CONSTRUCTION OF AN INDUSTRIAL SELECTION PERSONALITY TEST*

Walter V. Clarke Associates Inc., Providence, Rhode Island

WALTER V. CLARKE

A. INTRODUCTION AND RATIONALE

The resources available for measurement of the probable behavior of individuals in business and industry with respect to the area usually associated with temperament or personality, have been somewhat limited. On the other hand, the reporting of behavior through the medium of language depends on the fact that many words contain a definite and fairly clear-cut common core of meaning. Meanings of words, of course, do tend to vary somewhat, based on individual experiences of persons with particular words. Nevertheless, this condition is not such as to render language and communication a complete failure.

The successful use of self-describing words and phrases has been demonstrated widely in interest and personality inventories for counselling use. Denton (6) has demonstrated their usefulness in a forced choice personality test for a particular selection problem in industry. The present article reports research utilizing a somewhat different approach to the isolation and measurement of basic temperament traits in a wide variety of selection problems. Its end product was the test Activity Vector Analysis (AVA).

Therefore, in view of Thurstone's (13) findings of five basic clusters of adjectives descriptive of personality traits, together with related findings of other psychologists (1, 2, 3, 4, and 14), it was decided to attempt to measure the behavior of individuals through those words which best describe their behavior. It was decided to develop a kind of self-rating form without dependence upon face validity thus avoiding one of the usual weaknesses of inventories. It was decided to "infer the significance of responses from the groups differentiated" (11) in part, and also from the cluster analysis of response intercorrelations.

B. DEVELOPING INSTRUMENT

An original source list was prepared containing 183 words which were collected from standard dictionaries, dictionaries of synonyms and antonyms,

*Received in the Editorial Office on January 16, 1956, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

words used in describing personality and types in various textbooks, daily conversation, and reading. Subsequently, derogatory words were eliminated on the assumption that a semi-projective device such as was being sought would function best with manifestly derogatory terms absent. In addition, all words unsuitable for self-description were eliminated.

The resulting 81-word experimental form was administered to all applicants for jobs at a large machine manufacturing company employing up to 3,000 persons in Rhode Island. The individuals were asked to check those words: (a) anyone had ever used to describe them, (b) believed to be truly descriptive of them. A random sample of 100 was drawn from over 300 preliminary testings. No assumptions were made as to the meaning concerning whether or not a person checked a word under one of the above instructions or the other.

After tabulating the number of persons checking each word either under Instruction 1 or 2, the percentage of respondents was determined for each word on the list. Two words responded to by less than 5 per cent of the sample of 100 were eliminated. These were replaced and the list administered to another sample of 1,067 persons. The percentages of the sample responding to each word for both instructional conditions revealed no words checked by less than 5 per cent or more than 95 per cent of the sample.

In accordance with the trait clusters commonly found in personality analyses and the content of the factors reported by Thurstone (13) three groupings of words in the list of 81 were identified through study of tetrachoric word intercorrelations from the sample of 100 testees. These seemed to be clearly descriptive of the kind of behavior usually considered as: aggressive, sociable, and stable. These words served as the basis for three tentative scoring stencils using combined responses from the two response conditions. The sample of experimental forms were assigned scores for these three areas. All of the 81 words were correlated with these three scores, hereafter called Vectors. There appeared then to be the possibility of a fourth factor clustering around the word "fearful." All words were then correlated with a score obtained from a stencil based on words obviously related to the word "fearful." The results absorbed the words remaining on the form which had not proved to be assigned to the original three factors. The fourth factor content involved avoidant behavior as its theme.

C. ITEM WEIGHT STUDY

Weighted scores were derived from correlations of the 81 words with the cluster vector criterion (8, p. 273) and these correlated with unweighted scores in a sample of 100 cases. The result, $r = +.93 \pm .01$, indicates little

added accuracy of interpretation was to be gained by weighting. Therefore, item weighting was dropped from further study.

D. INITIAL NORMS

Preliminary norms were established using the first 100 cases selected in 1945. Later 500 cases were selected at random from 1,067 then tested and new norms established. *T*-Score Scales for each Vector and the total number of responses were set up with a mean standard score of 50 and 10 as one standard deviation. These scales were used to provide a profile which is interpreted in terms of factor content interactions and combinations.

E. ITEM RE-ANALYSIS

In 1947, 247 cases were randomly selected from the total population then tested and each word correlated tetrachorically with each of the four vector scores. Three words were eliminated from the list because of correlating less than $+.30$ with any vector score. An additional 75 words were correlated with the four vector scores and three substitute words selected. The final list of 81 words¹ meet all these standards: (a) None of them were responded to by more than 95 per cent or less than 5 per cent of individuals tested, and (b) all of them correlated $+.30$ or higher with at least one of the four conceptualized vectors of behavior.

Subsequently, to verify the original cluster selection and to examine the possibility of isolating additional clusters (in view of Thurstone having indicated seven areas of behavior and Cattell's studies pointing to an even larger number of primary traits), a study of the intercorrelations of 80 of the 81 words was accomplished² using IBM equipment.

From this analysis four clusters of words were isolated which agreed in general with the original selection. On the basis of these findings a new scoring stencil was derived which included some differential scoring of words under the two response conditions. Correlations between the newly determined vectors were either close to zero (positive) or were negative.³ Up to the present time only these four more or less basic vectors have been iso-

¹Copies of the test manual—showing the word list—may be secured from the author on request.

²Angle equivalents of the correlation coefficients were utilized to study cluster relationships using the technique reported by Cattell (5) and Guilford (7). An angle of 45 degrees projecting from the origin of the mathematical model of personality was set up as the limit around each cluster or vector. Each vector is the resultant of all the words making up the vector cone. Small sub-tables of intercorrelations were set up to facilitate the analysis and locate the basic constellations of words at 120° from each other.

³This indicates separations of clusters of approximately 90 degrees or more.

lated, although study of a larger selection of words might increase this number.

F. STANDARDIZATION

When over 15,000 tests had been accumulated, norms were established on 1,489 randomly selected cases. The total population consisted of all individuals who applied for work of whatever nature at three large manufacturing companies in Providence, Rhode Island.

In 1952 a randomly selected sample of 2,010 cases was used to establish T-Score Scales for responses under each of the two instructional sets: (a)

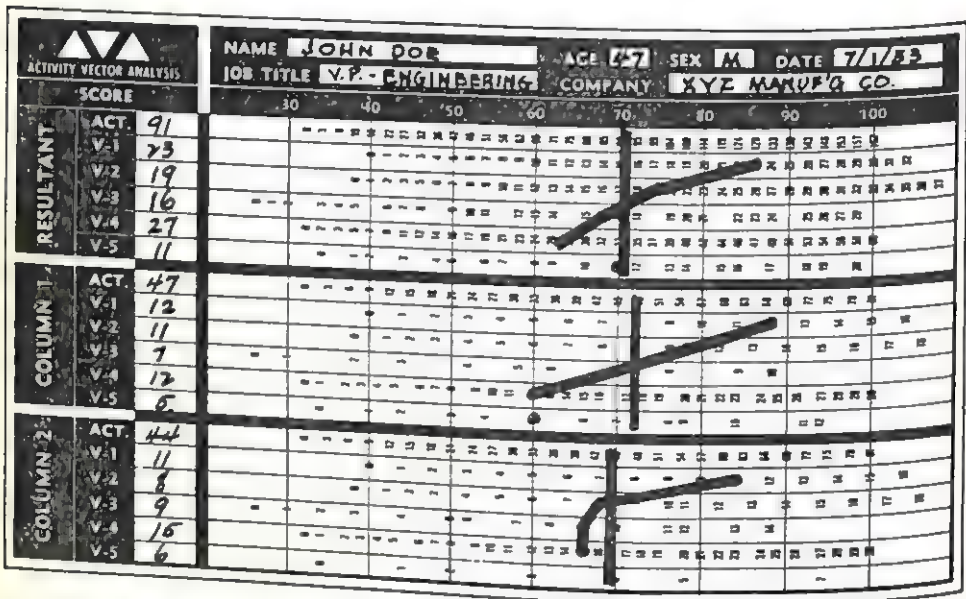


FIGURE 1
RECORD SUMMARY CARD

"Check those words anyone has used to describe you" (named Column 1);⁴ (b) "Check those words you believe actually describe you" (named Column 2).⁴ In addition, the scale for the combined responses (named Resultant)⁴ was further checked. This provides three profiles for interpretation as trait score syndromes two of which are based on separate homogeneous scaled summation of the two underlying response systems.

⁴See Figure 1.

G. RELIABILITY

1. *Split-half Reliability*

Split-half reliability for each of the Vectors and Activity (total number of responses) was computed separately for 100 cases selected at random from 5,000 tests available in 1948. Product-moment correlations for these data revealed the relationships presented in Table 1.

TABLE 1
COEFFICIENTS OF INTERNAL CONSISTENCY*
N = 100

Activity	$r = .97 \pm .02$
V-1	$r = .95 \pm .02$
V-2	$r = .94 \pm .02$
V-3	$r = .92 \pm .02$
V-4	$r = .95 \pm .02$

*Corrected by Spearman-Brown Formula.

2. *Test-Retest Reliability*

When considering the reliability of a temperament measurement it should be kept in mind that the analysis should be expected to reflect changes in behavior at different times and under different conditions. For test scores to remain constant when conditions affecting the subjects' behavior have altered would indicate a crude and relatively insensitive measuring instrument. However, the degree of sensitivity of the instrument will still depend upon the design and construction of the test.

Test-retest reliability was determined for 323 varied occupational cases with a mean retest interval of one year (minimum = one month) contributed by a large number of business and industrial companies located in various geographical areas of the United States. Product-moment correlations for the four basic Vectors and Activity (level of responsiveness) revealed the relationships indicated in Table 2.

TABLE 2
COEFFICIENTS OF STABILITY
N = 323

Activity	$r = .75 \pm .03$
V-1	$r = .73 \pm .03$
V-2	$r = .72 \pm .03$
V-3	$r = .62 \pm .03$
V-4	$r = .71 \pm .03$

The test-retest reliability of score profile shapes was determined for two occupationally different samples. The first involved 63 Tobacco Factory

Workers in Southeastern United States. The second sample included 57 managerial and technical employees submitted by a number of client companies. The results are shown in Table 3.

TABLE 3
TEST-RETEST RELIABILITY
(Profile Shapes)

	Resultant		Column 1		Column 2	
	Mean <i>r</i>	Med. <i>r</i>	Mean <i>r</i>	Med. <i>r</i>	Mean <i>r</i>	Med. <i>r</i>
a. Factory workers*	.71	.80	.67	.75	.69	.77
b. Managerial and technical**	.82	.88	.82	.92	.71	.78

*Interval—mean = 8.5 mos., min. = 3 mos. Sixty-three Tobacco Factory Workers.

** (Uniform 8.5 months' retest interval). Fifty-seven Managerial and Technical Employees.

The test-retest reliability coefficients appear to be highly satisfactory. They were higher, in fact, than had been expected because of the obvious limitations, so far as personality tests are concerned, which must be expected in test-retest results where no adequate controlling of the environmental situation is possible. The author judges that higher correlation coefficients for test-retest reliability with an instrument of the sort described in this paper would be undesirable, for the implication would be plain that such an instrument, especially a personality test, would be relatively insensitive to changes in the personality of the individual.

H. VALIDATION

1. *Similarity of Pattern*

Validity of the analysis would be indicated if it could be shown that significant differences exist between the personality patterns of persons successful on one kind of work, as compared to persons successful on work of a different nature. The profile cards were selected for all persons who could be classified under the following headings: President, Executive, Salesman, Accountant, Engineer, Teacher, Machine Operator. Obviously these titles represent a variety of occupations from a high level in the cultural hierarchy to a relatively low level. If this instrument is a useful tool for selecting salesmen, executives, accountants, or the like, there must be some significant differences among the patterns obtained from these people. The patterns were obtained by scoring the analyses for all persons successful on each of the jobs and computing the mean score and the standard deviation for each of the five scores involved in each score profile.

To determine that significant differences existed among these various patterns, critical ratios for the differences between the means were computed. Table 4 indicates the levels of confidence associated with the possibility of

TABLE 4
LEVELS OF CONFIDENCE WITH RESPECT TO POSSIBLE SIGNIFICANT DIFFERENCES BETWEEN
THE PROFILE SCORES OF AVERAGE PATTERNS

		President N = 30	Executive N = 27	Salesman N = 33	Accountant N = 20	Engineer N = 36	Teacher N = 30
Machine operator N = 78	Executive	Act. .225					
		V-1 .464					
		V-2 .216					
		V-3 .136					
		V-4 .363					
	Salesman	Act. .006	.017				
		V-1 .009	.021				
		V-2 .055	.007				
		V-3 .267	.325				
		V-4 .455	.408				
	Accountant	Act. .001*	.026	.081			
		V-1 .001	.001	.022			
		V-2 .001	.008	.001			
		V-3 .001	.001	.001			
		V-4 .001	.001	.001			
	Engineer	Act. .388	.109	.001	.192		
		V-1 .168	.220	.112	.001		
		V-2 .138	.021	.001	.003		
		V-3 .004	.109	.029	.001		
		V-4 .238	.005	.408	.001		
	Teacher	Act. .444	.001	.006	.113	.318	
		V-1 .001	.001	.001	.118	.001	
		V-2 .001	.041	.001	.125	.034	
		V-3 .001	.001	.001	.189	.001	
		V-4 .001	.001	.001	.448	.001	
	Machine operator N = 78	Act. .113	.004	.001	.000	.149	.068
		V-1 .001	.001	.201	.436	.001	.115
		V-2 .001	.105	.001	.248	.002	.229
		V-3 .001	.001	.001	.013	.001	.001
		V-4 .001	.001	.001	.016	.001	.001

*Italic figures indicate critical ratio of 3.0 or better, indicating that the differences are reliably greater than zero. This level of CR is chosen in view of the small although select samples involved.

actual differences of profile scores between occupations. It will be noted that fairly significant differences are found between most of the jobs and that where a significant difference is not found, there are certain similarities that where a significant difference is not found, there are certain similarities between the requirements of the jobs which will justify such a lack of difference. For example, although some difference does exist between "President" and the classification of "Executive" which included vice-presidents, general

managers, works managers, and the like, it should not be surprising if no truly significant differences were found, although there is some slight difference in the patterns as computed. However, significant differences do exist between presidents of companies and accountants, between presidents and teachers, etc. The fact that such significant differences do exist is an indication of the validity of the instrument for the purposes of selecting executives, salesmen, engineers, and the like.

As the reader will have undoubtedly inferred, the validation rationale here is essentially the same as was used by Strong. In his work, the development of scoring keys for his Vocational Interest blank was accomplished in accordance with the isolation of interest patterns characterizing successful incumbents of differentiable occupations.

A recent study of all ($N = 154$) "self-made presidents" for whom test profiles were available in our files, revealed that 141 (92 per cent) possessed profiles positively correlated with an "Administrative" profile (high vectors one and two, and low vectors three and four). This study supports the findings of an average pattern for successful Presidents and Executives in the above cited comparative study. Thirteen or 12 per cent of this self-made presidents group had profiles correlating zero or negatively with the previously established general profile. The mean of the correlations between the 154 obtained profiles and the standard profile was $+ .878$.

The validation of personality tests which are not interpreted in terms of, and do not yield a single numerical score, of which this is one, presents validation difficulties. This instrument is interpreted by integrating the scores in several areas of behavior. A man is not considered good sales material because of his score in any one area of behavior, but because of a particular combination and integration of scores from several areas of behavior measured by the test. This integration of scores of various behavior tendencies leads to a "personality profile" which may be found, empirically, to be characteristic of a "good salesman." The lack of a single numerical score forces the use of category judgments, or ratings, in determining the validity of the instrument.

An additional problem is encountered using personality tests in industrial situations in that the population cannot be selected in advance to any great degree. Therefore, the sample may or may not be representative of any single useful universe. This is to say that the universe is defined after the fact, in most cases. All that the psychologist can do is to provide as many samples as is practicable in the belief that this will make up for an admitted limitation.

In spite of the above limitations, the validity evidence as presently determined, suggests that this instrument has sufficient validity to serve as a valuable aid in the selection and placement of workers in business and industry. It is able to differentiate significantly between persons successful in such varied fields as executive, sales, accounting, machine operation, and teaching.

I. CORRELATION WITH OTHER TESTS

The intent of this research was to devise a measurement device for personality and especially for the measurement of human behavior in business and industry. The analysis was correlated with a variety of other tests used in industrial selection. Some of the resulting coefficients of correlation are shown in Table 5. Little or no relationship was found with these measures of intelligence or manual dexterity.

TABLE 5
INTERCORRELATIONS BETWEEN AVA RESULTANT SCORES AND OTHER TESTS

	N	Act.	V-1	V-2	V-3	V-4
Wonderlic	73	.15	.19	-.09	-.16	.24
Purdue Pegboard	73	-.22	-.09	-.16	-.15	.09
Otis S.A.	150	.11	.02	.12	.00	.00
Cardall	150	.08	.12	-.34	.08	.07

J. VECTOR AND VECTOR CRITERIA INTERCORRELATION

The Activity Vector Analysis of personality was constructed with the use of the vector criterion technique of item validation and selection in accordance with which the matrix of tetrachoric correlations among 80 of the 81 word items on the AVA form was used.⁵ Twenty of these words constituted the core words embodied in the vector criteria used in the construction of the scoring keys. The intercorrelations of these 20 core words together with the groupings which were used as criteria are shown in Table 6.

The zero order tetrachoric r 's in Table 6 were listed for each block of intercorrelations. These blocks of intercorrelations corresponded to the following test score combinations: Vector 1 versus Vectors 2, 3, and 4; Vector 2 versus Vectors 3 and 4; Vector 3 versus Vector 4; and the internal consistency of each of the four vectors (i.e., Vector 1 words versus other Vector 1 words, etc.).

⁵This master matrix has been deposited with the American Documentation Institute. Order Document No. 4781 from the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting in advance \$1.25 for microfilm or \$1.25 for photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

TABLE 6
INTERCORRELATIONS OF VECTOR CRITERION CORE WORDS ($N = 100$)

No.	Word	V-1					V-2						V-3				V-4			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	Nervy																			
2	Argumentative	.40																		
3	Opinionated	.30	.55																	
4	Stubborn	.42	.37	.37																
5	Dominant	.40	.50	.35	.09															
6	Personality Plus	.05	— .70	— .18	.06	.28														
7	Charming	.50	.20	.35	.19	.30	.75													
8	Appealing	.20	— .10	— .15	— .28	— .10	.42	.50												
9	Admirable	.40	.40	— .30	.26	.50	.40	.45	.55											
10	Smooth	.40	.05	— .05	.08	.70	.65	.30	.25	.35										
11	Attractive	.40	.00	— .40	.10	— .20	.52	.75	.60	.55	.25									
12	Magnetic	.60	— .10	.10	.07	.65	.23	.60	.20	.60	.50	.90								
13	Willing	— .10	— .15	.10	— .28	.20	— .33	— .60	— .30	— .20	.70	— .50	— .10							
14	Pleasant	.00	.00	— .85	— .14	.00	— .02	— .50	.35	— .20	.40	.00	— .60	.40						
15	Kind	.00	— .10	— .05	— .05	.05	— .08	— .60	.00	— .10	.20	.00	— .30	.30	.55					
16	Obliging	.00	— .10	— .10	— .05	— .15	— .32	— .40	.10	.05	.00	— .05	— .30	.50	.60	.35				
17	Harmonious	— .10	.10	.10	— .04	.20	— .26	.00	.00	.20	.20	— .25	.40	.30	.35	.10	.50			
18	God-Fearing	.00	.20	.45	.27	.10	.02	— .10	— .10	.10	.15	— .60	.20	.30	.00	.10	.20	.45		
19	Open-Minded	.00	.15	.10	— .15	.40	— .23	— .30	— .10	.15	.05	— .40	— .05	.00	.40	.20	.30	.45	.45	
20	Tolerant	— .05	.25	.40	.11	.00	— .35	.10	— .15	— .20	.15	— .05	— .20	.35	.20	— .10	.20	.55	.25	.50

V-1
Core
Words

V-2
Core
Words

V-3
Core
Words

V-4
Core
Words

Note: The correlations in the triangular cells are those which indicate the saturation relationships within the 4 vector criteria.

These lists of correlation coefficients were converted to Fisher Z Transformation values. The mean Z value and corresponding mean r was secured for each block in Table 6. The results are shown in Table 7.

TABLE 7
MEAN INTERCORRELATIONS OF VECTOR CRITERIA CORE WORD GROUPS, $N = 100$

Vectors	V-1	V-2	V-3	V-4	$N =$ Number of words
V-1	(+.38)				5
V-2	+.13	(+.52)			7
V-3	-.11	-.14	(+.46)		4
V-4	+.13	-.06	+.13	(+.45)	4
					<hr/> 20

Note: Correlations in parentheses show mean correlations of core words composing the respective vector criteria.

The intercorrelations of the Resultant Vectors of *AVA*, together with the reliability coefficients are reproduced in Table 8.

TABLE 8
INTERCORRELATIONS OF RESULTANT VECTOR SCORES, TEST-RETEST AND SPLIT-HALF RELIABILITY

Vectors	V-1	V-2	V-3	V-4	Maximum number of responses
V-1	(.73, .95)				32
V-2	+.15	(.72, .94)			37
V-3	-.53	-.31	(.62, .92)		28
V-4	-.62	-.43	-.10	(.71, .95)	60
					<hr/> 157

Note: Correlations in parentheses are Test-Retest reliabilities, $N = 323$, and corrected Split-Half reliabilities, $N = 100$, in that order. Vector intercorrelations are based on $N = 200$ unselected cases.

In order to facilitate the comparison of the correlations in Tables 7 and 8, the coefficients of Table 7 were adjusted in terms of equating for the number of possible score responses on the *AVA* Resultant Vectors. The formulae numbers 111 and 113 in Peters and Van Voorhis (2) were used for this purpose. The results are shown in Table 9.

The adjusted internal consistency coefficients of the criteria word groups in Table 9 are seen as high intermediate values between the two types of reliability coefficients of the vectors shown in Table 8. The original intercorrelations of words to be selected for core words as regards the intercorrelations of words to be selected for core words by cluster analysis involved a minimum of $r_{tet} = +.30$ for those words included in the criterion groups. In general, therefore, these mean intra-criteria correlations seem to be of an acceptable magnitude.

TABLE 9
 $N = 100$
 MEAN INTERCORRELATIONS OF VECTOR CRITERION CORE-WORD-GROUPS WHEN NUMBER
 OF WORDS IS SET EQUAL TO MAXIMUM NUMBER OF
 RESULTANT RESPONSES OF VECTORS

(a) Conversion factor	Vectors	V-1	V-2	V-3	V-4	N_1 = Num- ber of word responses
6.4	V-1	(+.80)				32
5.3	V-2	+.24	(+.85)			37
7.0	V-3	-.22	-.24	(+.86)		28
15.0	V-4	+.27	-.11	+.25	(+.92)	60

The correlations in Tables 7 and 9 clearly indicate the four word groups to be unrelated to each other. The inter-group correlations both positive and negative are of a low chance order. However, the within group or vector criterion mean correlations are consistently of a moderate and positive level of relationship.

In contrast with these relationships there are among the *AVD* vector correlations in Table 8, indications of the following negative relationships: Vector 1 versus Vectors 3 and 4, and Vector 2 versus Vectors 3 and 4 (although the V-2 versus V-3 relationship is relatively less strong).

In the construction of *AVD* each of the 81-word items was adopted for inclusion only after it had been demonstrated that it had an acceptably high correlation with at least one of the four vector criteria. Subsequently, scoring keys for the 81-word item form were built based on these item-vector criteria correlations. Therefore, an interpretation of the contrast between Tables 8 and 7 or 9 is offered with the aim of increasing understanding of the nature of this instrument and the theoretical concepts upon which it is based.

First of all, it seems apparent from the nature of the correlations involved that there remains substantial variance in each of the vector criteria word items which is unaccounted for by the variance which they appear to have in common. In terms of the semantics of the words in each criterion group and the test response-producing properties of these words, there are probably real differences in spite of the marked group similarity in these respects. Therefore, any larger number of word items all of which correlate significantly with a particular criterion group could do so on the basis of variance overlap with various portions of the total criterion variance; the common or non-common intra-criterion group variances. Theoretically, this might result in test score intercorrelations dissimilar to those among the test criteria.

Further, it is essential in the evaluation of the significance of obtained

vector intercorrelations (as well as the vector criteria intercorrelations) to recognize the personality theory and psychometry underlying the vector concept and the corresponding vector measures. The traits postulated and attempted to be measured are unipolar in nature, each with magnitudes of measurable manifestation or attribute from a theoretical minimum to a theoretical maximum. Personality, therefore, is conceptualized as a sphere wherein a number of tendencies to behave in differing strengths, interact or influence each other so as to result in a complex compromise direction or vector of behavioral tendency. Each of the *AVA* related trait scores were named vectors on the basis of this conceptualization. In terms of word intercorrelations the location of each Vector in space is the resultant of the magnitudes and positions of the words which produce it (within, not merely upon the theoretical sphere). Therefore, the intercorrelations of the vectors represent the spatial relationships among them, i.e., a zero correlation representing a 90° separation and a minus correlation representing greater separation. It would seem then that the transition from vector criteria to test vectors (Tables 7 and 8) represents an increase in spatial separation among these four measures over that possessed by the criteria. That each of the four vectors is spatially well removed from the other three can be seen in Figure 2.

K. PROFILE INTERPRETATION

When the verbalized interpretations of these vectors were made, however, although there are some words present with apparently opposite meaning, the basic concept was found to be not opposite, but different. The importance of these negative relationships became evident only after a considerable number of cases were studied.

Most people tend to have a balance of personality traits, such, that as aggressiveness increases, independence also increases. When this is not the case and aggressiveness is matched by dependence, a state of vacillation and indecision is created, and further deviation is suggestive of some areas of non-conformity with the law. Since our population did not have an adequate sampling of such deviant individuals, the high negative correlations may not be truly indicative of the relationships between the vectors. A wider sampling of all personality types in the normal population will be obtained for further study in this area.

It was necessary, if the above concepts were to be applied empirically, to establish some method of scaling words along the continuum, as well as determining to what extent each word pointed out the direction of the basic

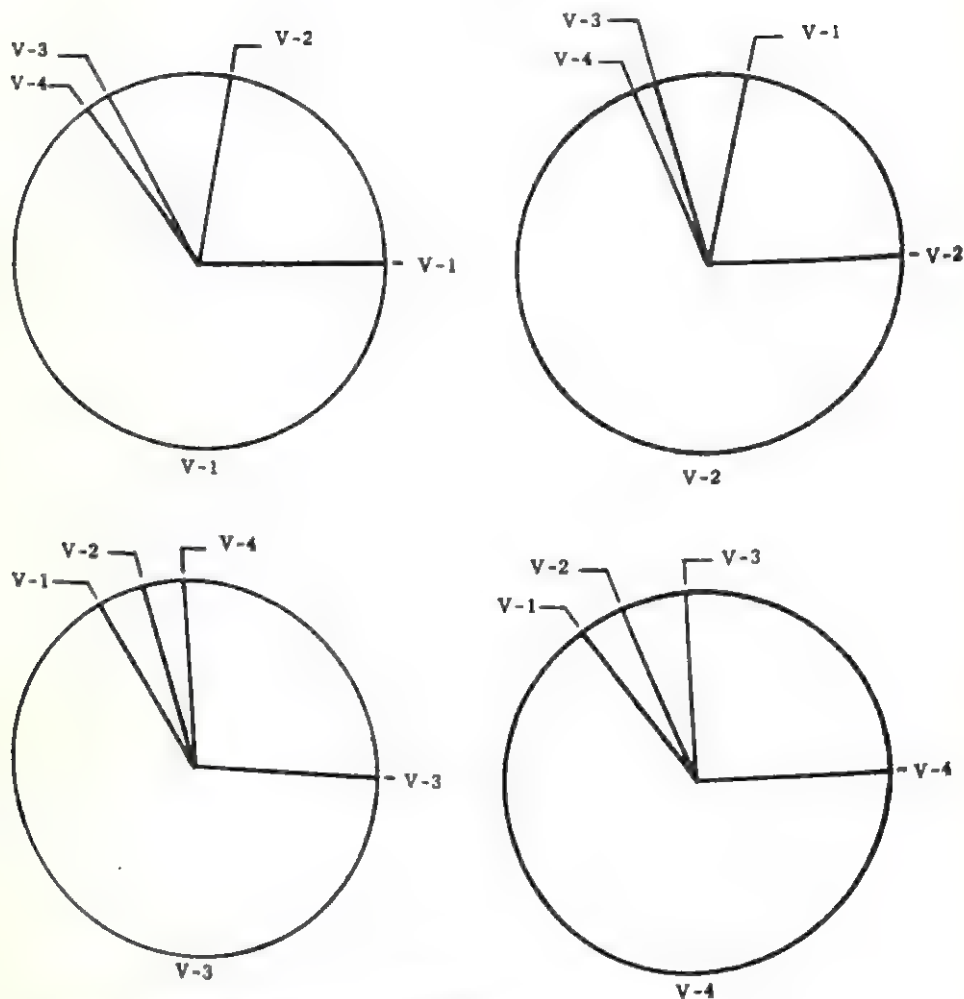


FIGURE 2
VECTOR INTERCORRELATIONS REPRESENTED IN TERMS OF ANGULAR RELATIONSHIPS

vector. Since no perfect correlation existed with the central concept, the directions created a cone with a large number of word continuums within it, all of which clustered around the central basic vector concept.

The exact location of any word continuum is not based on the correlation with the vector concept, but by its relationship to all vector concepts, since no word was found that was not affected by at least three vectors, each of which obviously affected its position in space.

Having determined these relationships, it was necessary to determine the location of the word along its continuum with relationship to the origin.

It was assumed that the social acceptability of a word is a measure of its intensity. From the total number of times a word was checked, the standard

deviation was computed on the additional assumption that the distribution of words form a normal curve. In this way the location of any word may be scaled in terms of its sigma.

All words in the experimental studies were now sorted by vector concept as well as by direction within the vector concept and laid out along their continuum in accordance with their sigma location. This rather complex scale was then combined into one vector scale for convenience. To make the scale more complete for practical purposes, additional words were added. Although statistical determination of the value of each of these words was not made, it is believed that their locations are fairly accurate, since the analysis had indicated rather clear-cut differences in the concepts of the continuums. For example, the "bold, nerry, egotistical" continuum contained more antagonism than the "courageous, go-getter, persistent" continuum. This, together with the scales originally established, served as a fairly accurate guide for interpolation. These scales were then assembled together to serve as a guide in making interpretations.

L. SEX DIFFERENCES

It may be noted here that no significant sex differences have been found in the use of Activity Vector Analysis. This conclusion is based on a study of 161 cases (74 male, 87 female) selected at random from 300 cases available in 1946. Tests of the differences between the means of the male and female groups for activity and for the four vectors showed none as significant.

M. SUMMARY

An adjective and adjective phrase self-rating form consisting of 183 items, sampling many different aspects of temperament was assembled. Eighty-one of these were selected as non-derogatory and seemingly suitable for study. These were subjected to a frequency response pre-test with a sample of 100 industrial applicants. After replacing two words rejected for too limited response value, word intercorrelations and a cluster type factorial study was accomplished using another industrial sample of 1,067 cases. This yielded three "factors": aggressiveness, sociability, and emotional stability. A fourth factor was derived from a scoring key based on the word "fearful" which seemed to be associated with a fourth but previously inadequately clarified cluster. A later study of 227 cases produced verification and some refinements. The fourth factor appears as social adaptability or dependence.

T-Score Scales were established for these four measures in addition to total number of test responses or Activity as a fifth measure. The original standardization was based on 500 randomly selected industrial applicants. The

later revisions of the scales are based on 1,489 cases and 2,010 cases selected at random from a test population of over 15,000 industrial applicants.

A number of validities have been investigated. Significant differences have been found between score profiles of varied occupations. Although these validity evidences may not be conclusive, they are very promising.

N. CONCLUSIONS

1. A new analysis of temperament has been developed which appears very appropriate for use in business and industry.
2. The instrument appears to have very satisfactory internal consistency and test-retest reliability.
3. This test or analysis appears to have satisfactory validity for occupational differentiation.
4. Further research is clearly indicated with respect to a variety of additional job success validities as well as interpretation validity.⁶

REFERENCES

1. BURT, C. L. The analysis of temperament. *Brit. J. Med. Psychol.*, 1938, 17, 155-188.
2. ———. The factorial study of emotions. In M. L. Reymert, Ed., *Feelings and Emotions*. New York: McGraw-Hill, 1950. (Pp. 531-551.)
3. CATTELL, R. B. The principal trait clusters for describing personality. *Psychol. Bull.*, 1945, 42, 129-161.
4. ———. Description and Measurement of Personality. New York: World Book, 1946. Pp. 602.
5. ———. Factor Analysis. New York: Harper, 1952. Pp. 462.
6. DENTON, J. C. Building a forced-choice personality test. *Pers. Psychol.*, 1954, 7, 449-459.
7. GUILFORD, J. P. Psychometric Methods. New York: McGraw-Hill, 1936. Pp. 566.
8. HULL, C. L. Aptitude Testing. New York: World Book, 1928. Pp. 535.
9. JENKINS, T. N. Some contributions in support of a neutral theory of personality. *Trans. New York Acad. Scis.*, 1950, 13, 9-12.
10. ———. The neutral theory of personality: new evidence and a review of the argument. *Trans. New York Acad. Scis.*, 1955, 17, 315-330.
11. MEEHL, P. E. An investigation of a general normality or control factor in personality testing. *Psychol. Monog.*, 1945, 59, No. 4.
12. PETERS, C. C., & VAN VOORHIS, M. A. Statistical Procedures and Their Mathematical Bases. New York: McGraw-Hill, 1940. Pp. 478.
13. THURSTONE, L. L. The vectors of the mind. *Psychol. Rev.*, 1934, 41, 1-32.
14. ———. The dimensions of temperament. *Psychometrika*, 1951, 16, 11-20.

Walter V. Clarke Associates, Inc.
500 Angell Street
Providence 6, Rhode Island

⁶A number of such studies are either in progress or ready to be reported at the time this paper is submitted.

A SUGGESTED MEASURE OF BIVARIATE RELATIONSHIP BASED ON THE ELLIPSE OF CONCENTRATION*

Humm Personnel Consultants, Los Angeles

DONCASTER G. HUMM

For certain computational and interpretative purposes in the field of psychometrics, a measure of bivariate relationship and its alienation coefficient are offered for the critical consideration of statisticians.

The measure is designated as Upsilon (Υ) and its alienation coefficient as Lambda (Λ). For simplicity, consideration has been limited to standard scores, where the standard score is:

$$z = (X - \bar{X})\sigma^{-1}$$

In this formula, X is any score; \bar{X} , the mean; and σ the standard deviation. Application of the measures can be extended to raw scores by use of suitable formulas, but these are not discussed in this paper.

These measures have been sought because of certain inconveniences in the use of Pearson's product-moment correlation (r): (a) Pearson's r is difficult to interpret, since it approaches ± 1 by unequal increments. (b) These unequal increments result in a rather wide range of sampling values of r and N , within which normal curve approximations will not be adequate (6, p. 158; 3, pp. 215-221; 5, p. 155). (c) The x - and y -axes do not report equivalent deviations from each of the two regression lines, since $\bar{z}_y = rz_x$ assumes error in \bar{z}_y only; and $\bar{z}_x = rz_y$ assumes error in \bar{z}_x only. (d) It is difficult to determine critically the presence of curvilinearity precluding the employment of r and the correction for such curvilinearity.

Upsilon and Lambda are based on the family of homothetic ellipses known to delimit various degrees of concentration of the correlation surface. Considerable work has already been done on the study of this parameter family.

Cramér (1) has given a formula which, reduced to z -scores, is:

$$z^2_x - 2rz_xz_y + z^2_y = c^2 \quad (1)$$

in which r is Pearson's product-moment coefficient of correlation and c is any degree of concentration. Cramér calls this the "Ellipse of Inertia."

*Received in the Editorial Office on January 25, 1956, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

When the first and second moments of the variables coincide with those of one of these ellipses, since then $z_x = z_y = 1$, the equation becomes:

$$\begin{aligned} c &= \sqrt{2(1-r)} \text{ for } r > 0, \text{ and} \\ c &= \sqrt{2[1-(-r)]} \text{ or } \sqrt{2(1+r)} \text{ for } r < 0 \end{aligned} \quad (2)$$

This is the equation for the standard deviation of the difference (σ_d) for $r > 0$; and the standard deviation of the sum for $r < 0$. Cramér calls this the "Ellipse of Concentration." He notes that the axes of this ellipse are orthogonal mean square regression lines rotated 45° from the z_x and z_y axes.

The equation for these axes affords a simplification of Cramér's equation.¹ For $r > 0$, let M be the major axis and m the minor axis. Then

$$z_x = \sqrt{2}(M - m) \text{ and } z_y = \sqrt{2}(M + m) \text{ or} \quad (3)$$

$$M = \sqrt{2}(z_x + z_y) \text{ and } m = \sqrt{2}(z_x - z_y) \quad (4)$$

in which it is easy to determine that

$$\frac{M^2}{4(1+r)} + \frac{m^2}{4(1-r)} = 1 \quad (5)$$

Since Equation 2 is a formula for σ_d when $r > 0$ and for σ_s when $r < 0$, and since (from Equation 4) M is measured in sums of z -scores and m in differences of z -scores; it follows that the Ellipse of Concentration is reportable in terms of σ_s and σ_d , where M is the locus of *zero* differences and rows of cells parallel to M the loci of *equal* differences, while m is the locus of *zero* sums and rows of cells parallel to m the loci of *equal* sums, with orthogonal measurement from the respective axes.

As Treloar (6, p. 159) points out, the dimensional ratio of the major to the minor axis for the normal surface ellipse, plotted in terms of relative deviates (or z -scores), is

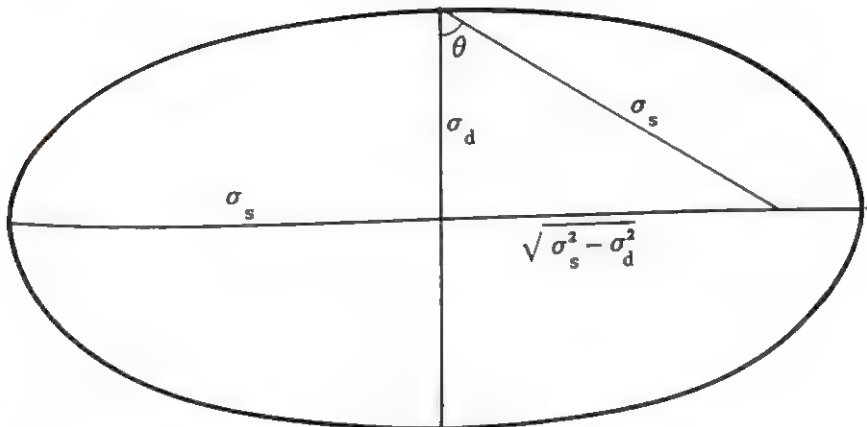
$$\sqrt{\frac{1+r}{1-r}},$$

(a ratio which is shown in Table 1 as a function of Y and Λ). This ratio is very intimately related to intensity of association. Fisher (3) suggests replacing r by the natural logarithm of this ratio for "certain interpretative and computational purposes" where

$$\operatorname{arctanh} r = z' = \frac{1}{2}[\log_e(1+r) - \log_e(1-r)] \quad (6)$$

in which $-\infty \leq z' \leq +\infty$.

¹Equations 3 and 4 and their relationship were kindly pointed out by Professor William Smith of Bucknell University. He refers to Cramér (1, p. 284).



$$\sin \theta = \frac{\sqrt{\sigma_s^2 - \sigma_d^2}}{\sigma_s} \quad \cos \theta = \frac{\sigma_d}{\sigma_s} \quad \text{vers } \theta = 1 - \frac{\sigma_d}{\sigma_s}$$

FIGURE 1
 σ_s AND σ_d AS IN POSITIVE BIVARIATE FUNCTIONS

Fisher's equation is, in effect, a rectifying of the curvilinearity with which Pearson's r approaches perfect correlation. His z' reduces the coefficient of alienation, where $k = \sqrt{1 - r^2}$, to almost invariantly equal intervals.

Our study of the Ellipses of Inertia attempts to simplify the consideration of bivariate relationship by rotation of the cartesian axes as is shown in Equations 3 and 4. In considering the ellipse as a conic section, since the theoretical upper limit of both x and y is infinity, we were obliged to consider a cone with its apex an "ideal" point—in other words, a cylinder. In this case, when the angle of cut is orthogonal to the axis of the cylinder (or is zero), the ellipse of inertia becomes a circle. This happens when the bivariate relationship is zero. When the angle of cut reaches 90° from the horizontal, the ellipse becomes a straight line, as it does when the relationship is perfect.

It has several times been demonstrated in the literature that the angle of cut is equal to the angle ordinarily used to measure eccentricity. This is established because the minor axis is the radius of the cylinder; and the major axis, by reason of being also the distance from a focus of the ellipse to the end of the minor axis, is the hypotenuse of both the right angle of which θ is an acute angle and the right triangle of which the angle of cut is an acute angle. Thus the two right triangles have a leg and a hypotenuse equal and are themselves equal. As a consequence, we considered functions of Theta in arriving at Upsilon and Lambda.

In ordinary practice the eccentricity of an ellipse is measured by $\sin \theta$. However,

$$\sin^2 \theta + \cos^2 \theta = 1 \quad (7)$$

It is obvious that both r with k and $\sin \theta$ with $\cos \theta$ are in the form of the equation for a circle, and it is also obvious that $\sin \theta$ approaches perfect correlation, in the same fashion as r does, since $r^2 + k^2 = 1$.

However, as previously noted, the standard deviation of the difference, σ_d , is a measure of departure from perfect correlation in positive relationship and the standard deviation of the sum, σ_s , in negative. In the ellipse of concentration these are respectively dependent on each other; so that the alienation is expressible:

$$\pm \Lambda = \left(\frac{\sigma_d}{\sigma_s} \right)^{\pm 1} = \cos \pm \theta \text{ or } +\Lambda = \frac{\sigma_d}{\sigma_s} \text{ and } -\Lambda = \frac{\sigma_s}{\sigma_d} \quad (8)$$

The \pm signs are necessary, since σ_s is the major axis for positive relationship and σ_d for negative. Note that for the Ellipse of Concentration $\sigma_s/\sqrt{2}$ is one axis and $\sigma_d/\sqrt{2}$ the other, but that in the equation for Λ these $\sqrt{2}$'s cancel out.

Since the versine of an angle is equal to one minus the cosine, letting $Y = \text{vers } \theta$

$$\pm Y = 1 - \left(\frac{\sigma_d}{\sigma_s} \right)^{\pm 1} \text{ or } +Y = 1 - \frac{\sigma_d}{\sigma_s} \text{ and } -Y = 1 - \frac{\sigma_s}{\sigma_d} \quad (9)$$

$$\text{and } \pm Y + \pm \Lambda = 1 \quad (10)$$

While it is possible to state these equations in general terms by the use of vector notation, the treatment we have used seems to follow standard geometric practice. Equation 8 is in a plus and minus form for the reason that a measure orthogonal to another measure is expressible as minus the reciprocal of that other measure.

Since Equation 10 is in the form of an equation for a straight line, Upsilon approaches perfect correlation by equal increments.

The arrays of differences of z -scores for any sum of z -scores and vice versa are measurable by an adaptation of a theorem in geometry which states that the square of either semi-axis is to the square of the other as the rectangle of the former is to the square of the ordinate. Thus,

$$M^2 : m^2 :: y(2M - y) : a^2 \quad (11)$$

in which y is the distance from the ordinate, a , along M to curve of ellipse, so that $x + y = M$. Substituting $y = M - x$,

$$\frac{m^2}{M^2} = \frac{a^2}{M^2 - x^2} \quad (12)$$

Since $m/M = \Lambda$, inasmuch as the Ellipses of Inertia are homothetic, the axes of all are in the same proportion.

$$a_1 = \Lambda \sqrt{M^2 - x_1^2} \quad (13)$$

for ordinates from M -axis, and similarly

$$a_2 = \Lambda \sqrt{m^2 - x_2^2} \quad (14)$$

for ordinates from the m -axis. Hence the limit of the axes is

$$m = +\Lambda M \text{ and } M = -\Lambda m \quad (15)$$

The theoretical range is, of course, infinity; but sampling range depends on the number of cases, N , as Kelley (4) points out, in the equation:

$$N(1 - \alpha) = .5 \quad (16)$$

in which $\alpha = 1 - 2q$, where q is a tail of a distribution and $p + q = 1$. E.g., for 1,075 cases the range $= \pm 3.5 \sigma$.

When these finite values are substituted for the range, the lengths of the major axis and the minor axis of the ellipse covering the range of sampling are respectively as follows:

$$M = \frac{2n}{\sqrt{1 + \Lambda^2}} \quad (17)$$

$$m = \frac{2n\Lambda}{\sqrt{1 + \Lambda^2}} \text{ and vice versa} \quad (18)$$

and

in which n is the number of σ 's in each variant. When the distributions are both approximately normal but n 's slightly different, an approximation may be made:

$$\bar{n}_0 = \sqrt{n_x n_y} \quad (19)$$

Since the lack of normality of the sampling distribution of r has been demonstrated by Fisher (loc. cit.) as being due to unequal increments and as being correctable by use of his z' , which is based also on elliptical considerations, we suggest that the standard error of Y is

$$\sigma_r = \frac{1}{\sqrt{N-3}} \quad (20)$$

We offer this for further investigation.

If further study confirms the validity of Y and Λ as measures, the following uses of them are suggested:

1. *They will be useful for the linear interpretation of Pearson's product-moment correlation coefficient.*

$$\text{Since} \quad \sigma_s^2 + \sigma_d^2 = 4 \quad (21)$$

$$\text{and} \quad r = \frac{\sigma_s^2 - \sigma_d^2}{4} \quad (22)$$

$$\text{and} \quad k = \frac{\sigma_s \sigma_d}{2} \text{ in which } k = \sqrt{1 - r^2} \quad (23)$$

$$\text{and} \quad z' = \tanh^{-1} r \quad (24)$$

it follows that r , k , z' , Y and Λ are functions of each other, as shown in Table 1.

Table 1
The Relationships among Bivariate Measures

	$\frac{+T}{1 - \frac{\sigma_d}{\sigma_s}}$	$\frac{+\Lambda}{1 - \Lambda}$	$\frac{r}{1 - \sqrt{\frac{1-r}{1+r}}}$	$\frac{k}{1 - \frac{k}{1 + \sqrt{1+k^2}}}$	$\frac{z'}{1 - e^{-z'}}$
$+T$					
$+\Lambda$	$1 - T$	$\frac{\sigma_d}{\sigma_s}$	$\sqrt{\frac{1-r}{1+r}}$	$\frac{k}{1 + \sqrt{1+k^2}}$	$e^{-z'}$
r	$\frac{1 - (1-T)^2}{1 + (1-T)^2}$	$\frac{1 - \Lambda^2}{1 + \Lambda^2}$	$\frac{\sigma_s^2 - \sigma_d^2}{4}$	$\sqrt{1 - k^2}$	$\tanh z'$
k	$\frac{2(1-T)}{1 + (1-T)^2}$	$\frac{2\Lambda}{1 + \Lambda^2}$	$\sqrt{1 - r^2}$	$\frac{\sigma_s \sigma_d}{2}$	$\sqrt{1 - \tanh^2 z'}$

(For $-T$ and $-\Lambda$, the formulas must be adjusted.)

$$\text{Since} \quad -1 \leq r \leq +1 \quad (25)$$

$$\text{and} \quad -1 \leq Y \leq +1 \quad (26)$$

interpretation is possible in closed intervals within the same upper and lower bounds.

In Table 2 the relationships among these measures are given in tabular form. (The author has also prepared a more detailed table, available on request, of the relationship of r and Y . It is too space-consuming to be presented here.)

2. *The use of arrays of sums and differences of z-scores is also suggested for studies of and corrections for curvilinearity in the orthogonal mean square*

TABLE 2
RELATIONSHIPS AMONG BIVARIATE FUNCTIONS

T	A	r	k	z'
1.00	.00	1.000	.000	∞
.95	.05	.995	.100	2.994
.90	.10	.980	.199	2.298
.85	.15	.956	.293	1.897
.80	.20	.923	.385	1.609
.75	.25	.882	.471	1.385
.70	.30	.834	.552	1.201
.65	.35	.782	.623	1.050
.60	.40	.724	.690	.916
.55	.45	.663	.749	.798
.50	.50	.600	.800	.693
.45	.55	.536	.844	.599
.40	.60	.471	.882	.511
.35	.65	.402	.914	.431
.30	.70	.342	.940	.356
.25	.75	.280	.960	.288
.20	.80	.220	.976	.224
.15	.85	.161	.987	.162
.10	.90	.105	.994	.105
.05	.95	.051	.999	.051
.00	1.00	.000	1.000	.000

regression lines. These studies are essential, since Y and A are both calculated on the assumption that the data are normally distributed.

Since, when the bivariate surface is normal, the M -axis is the locus of zero differences and the m -axis of zero sums, and the sums and differences are symmetrically distributed; then for any array of differences for any given sum, and for any array of sums for any given difference, it follows that the mean of each array should be zero. Hence any lack of linearity may be ascertained by the following formula:

$$\frac{-Mn_a}{\sigma_{mna}} = t \quad (27)$$

where Mn is the mean of the array and t a critical ratio.

If $t \geq 1.96$, it is shown that at least a 95 per cent of confidence exists that the array is skewed; if $t \geq 2.58$, at least a 99 per cent is afforded.

Curvilinearity can be determined by tabulating differences in z -scores ($z_x - z_y$) against sums of z -scores ($z_x + z_y$) on columns and rows (similar to a correlation scatter chart), using the abscissa for the positive axis (M -axis) and the ordinate for the negative axis (m -axis), as shown in Figure 2. This figure gives an example of curvilinearity as reported by Davis

(2, pp. 42 and 51). The m -axis has a form similar to $y = \tanh x$, and the M -axis similar to $y = a + bx + \delta$ where

$$\delta = \frac{a}{e^{n(X-x)} + e^{n(x-X)}} \quad (28)$$

in which X is the value of x corresponding to the maximum deviation of y from a straight line.

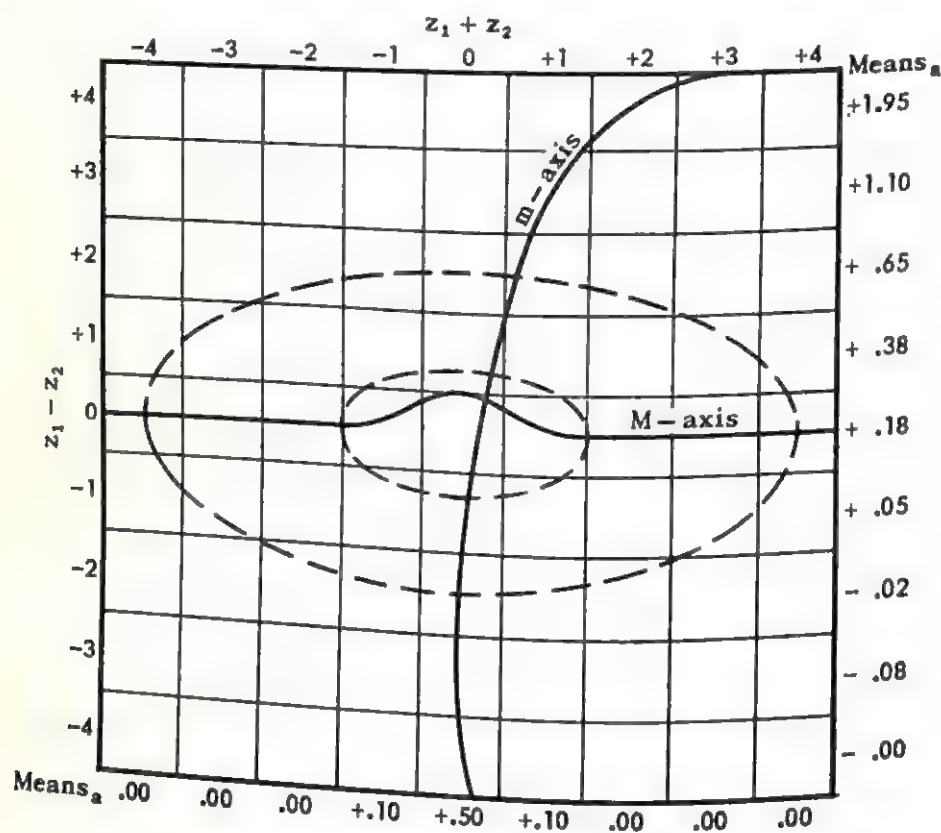


FIGURE 2
SCHEMATIC DISTRIBUTION CHART

In Figure 2 the limits of the ellipse of concentration are, by definition, the inner ellipse; while the ellipse defining the range of sampling is the outer ellipse.

Davis has given a number of ways in which such curvilinearity may be rectified. This is one of the necessary steps in the calculation of a mathematically valid Upsilon. Another is the employment of Kelley's formula (For-

mula 16) to determine whether or not the data cover the normal range for the number of cases.

3. *The use of sums and differences of z-scores and σ_s and σ_d makes it possible to calculate Pearson's r without extra calculation and to determine whether or not the bivariate relationships are such that Y and r may be validly used.*

Inasmuch as both r and Y are predicated on the analysis of normal distributions, it is necessary that the bivariate scatter-field be normally distributed in order that either of these will measure validly in all respects. The necessary conditions are (a) That the regression be linear. Formula 27 determines this in terms of level of confidence. (b) That the scatter-field be symmetrical (homoscedastic and homoclitic): The analysis of the arrays of sums and of differences of z-scores, as suggested under 2, will determine this. (c) That the sampling range be in accord with normal expectancy. Formula 16 will determine this. It should be noted that this third condition is not essential when it is understood that the relationship in a limited sampling is being determined.

4. *Upsilon and Lambda are not offered at this time as substitutes for r and k , but rather as interpretations of these latter two measures and as means of determining conditions precedent for the use of all four, and then only after confirmatory investigation by other students.* It should be pointed out that this paper merely discusses the means of determining simple bivariate relationship, and does not touch upon partial and multiple correlation. If Upsilon and Lambda prove to be valid measures, much more investigation must be carried on before they may be substituted for other techniques.

REFERENCES

1. CRAMÉR, H. *Mathematical Methods of Statistics*. Princeton University Press, 1946.
2. DAVIS, D. S. *Empirical Equations and Nomography*. McGraw-Hill Book Co., Inc., New York, 1943.
3. FISHER, R. A. *Statistical Methods for Research Workers*. Olive & Boyd. London, 1925.
4. KELLEY, T. L. *Statistical Method*. Macmillan Company. New York, 1924.
5. PETERS, C. C., & VANVOORHIS, W. R. *Statistical Procedures and Their Mathematical Bases*. McGraw-Hill Book Co., Inc. New York, 1940.
6. TRELOAR, A. E. *Elements of Statistical Reasoning*. John Wiley & Sons, London, 1939.

*Humm Personnel Consultants
1219 West 12th Street
Los Angeles 15, California*

THE PERSONALITY PROFILES OF LOAN OFFICE MANAGERS*

Walter V. Clarke Associates, Inc.

WALTER V. CLARKE

A. INTRODUCTION

This study was made to test the skill of a client company president and Activity Vector Analyst in his use of Activity Vector Analysis (*AVA*) profile selection standards in the employment of potential branch managers for personal loan offices. Subsequent to completion of the *AVA* Program of Personnel and Human Relations Training, the Analyst returned to his business and evaluated the temperament requirements of the Loan Office Manager's (D.O.T.: 0-85.10) and Assistant Loan Office Manager's (D.O.T.: 0-98-08) jobs in his organization in terms of the *AVA* profile implications thereof. The results of his evaluations form the selection standard hypothesis of this study.

B. PROBLEM

To determine the relevance of selected *AVA* score profiles for prediction of success as a Loan Office Manager or Assistant Manager.

C. JOB STUDIED

Loan Office Manager (and Assistant Manager) in branch office of a chain of personal loan establishments (D.O.T.: 0-85.10 and 0-98.08). Directs the management of loan agency; supervises employees and assigns duties; plans and administers methods of procedure; may hire and fire employees; investigates financial standing and reputation of prospective customer who is seeking loan; approves or disapproves loans; supervises collection of bad accounts; makes up collection notes to customers; keeps records of collections.

D. GEOGRAPHICAL LOCATION

Small mill towns in Piedmont Region of the South; small farms, cotton mills, and small manufacturing. In most towns colored population is 25 per cent or less. Except for Ashton and Atlanta, essentially small mill towns in character.

*Received in the Editorial Office on February 1, 1956, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

E. THE CRITERION

A composite Satisfactory-Unsatisfactory rating of each individual was based on the following:

TABLE 1
TETRACHORIC INTERCORRELATIONS OF COMPOSITE AND SUBORDINATE CRITERION RATINGS

	I	II	III	IV	V	N
I Supervisor's Rating		+.86	+.96	+.99	+1.00	89
II Delinquency Rating			+.64	+.72	+.86	50
III Profit Rating				+.77	+.96	50
IV Growth & Improvement Rating					+.99	89
V Composite Rating						89

1. District Supervisor's dichotomous and subjective over-all rating.
2. District Supervisor's dichotomous rating of the man's ability to keep his accounts paying off their loans (delinquency rating).¹
3. District Supervisor's dichotomous rating of the man's ability to produce a standard profit in his office each month (profit rating).¹
4. Combined District Supervisor and company president's dichotomous rating of the man's growth and improvement.

F. CHARACTERISTICS OF SAMPLE ($N = 89$)

1. Sub groups: (a) Office Managers—hired before *AVA* was used; $N = 34$. (b) Assistant Managers—hired before *AVA* was used; $N = 26$. (c) Assistant Managers—hired with *AVA*; $N = 29$.
2. Experience: Shown in Table 2.
3. Age: Shown in Tables 3 and 4.

TABLE 2
EXPERIENCE CHARACTERISTICS OF SATISFACTORY AND UNSATISFACTORY RATED LOAN OFFICE PERSONNEL

Sub Groups	(N)	Mean—Years (Mos.)		Range—Years (Mos.)	
		Satisfact.	Unsatisfact.	Satisfact.	Unsatisfact.
*a. Office Mgrs.	(34)	3.22 yrs.	1.77 yrs.	2-8 yrs.	1-4 yrs.
*b. Asst. Mgrs.	(26)	20.92 mos.	13.43 mos.	9-54 mos.	4-30 mos.
*c. Asst. Mgrs.	(28)	9.17 mos.	6.20 mos.	1-54 mos.	1-24 mos.

*Difference between means significant; $P = .006$, $t = 2.76$.

**Difference between means not significant; $P = .11$, $t = 1.60$.

***Difference between means not significant; $P = .46$, $t = .73$.
(One case for which experience data unavailable)

¹Restricted to Office Managers at time of study.

TABLE 3

AGE CHARACTERISTICS OF SATISFACTORY AND UNSATISFACTORY RATED OFFICE MANAGERS

		Mean	SD	Range
Sub Group (a) Total	$N = 34$	27.7 yrs.	5.3 yrs.	21-42 yrs.
Satisfactory Rated	$N = 23$	26.1 yrs.	3.3 yrs.	21-32 yrs.
Unsatisfactory Rated	$N = 11$	31.1 yrs.	6.9 yrs.	21-42 yrs.

$$r_{bis} (\text{age vs. success}) = -.58; SE r_{bis} = \pm .13.$$

TABLE 4

AGE CHARACTERISTICS OF SATISFACTORY AND UNSATISFACTORY RATED ASSISTANT OFFICE MANAGERS

	Satisfactory	Unsatisfactory
Sub Group (b) mean	25.2 yrs.	25.9 yrs.
Asst. Mgrs., $N = 26$ SD	2.8 yrs.	3.2 yrs.
*Sub Group (c) mean	26.1 yrs.	25.4 yrs.
Asst. Mgrs., $N = 29$ SD	3.7 yrs.	3.4 yrs.

$$**t \text{ test of difference between means} = .496.$$

TABLE 5

EDUCATIONAL LEVEL VERSUS PERFORMANCE RATING OF OFFICE MANAGERS

	Perform. rating		Total
	Unsatisfact.	Satisfact.	
(a) Mgrs. ($N = 34$)			10
High school and more	1	9	24
High school only	10	14	34
Total	11	23	

$$r_{tet} (\text{education vs. Rating}) = +.555.$$

$$\chi^2 = 3.23; P = .06 \text{ (With Yate's Correction, } \chi^2 = 1.95, P = .17); (F_E \text{ too small}).$$

TABLE 6

EDUCATIONAL LEVEL VERSUS PERFORMANCE RATING OF ASSISTANT OFFICE MANAGERS HIRED BEFORE AVA

	Perform. rating		Total
	Unsatisfactory	Satisfactory	
(b) Assistant Mgrs. ($N = 26$)			8
High school and more	4	4	18
High school only	12	6	26
Total	16	10	

$$r_{tet} = +.25 \text{ (not significant).}$$

$$\chi^2 = .137.$$

TABLE 7
EDUCATIONAL LEVEL VERSUS PERFORMANCE RATING OF ASSISTANT OFFICE MANAGERS
HIRED AFTER *AVA* WAS ADOPTED

(c) Assistant Mgrs. (<i>N</i> = 28)	Perform. rating		Total
	Unsatisfactory	Satisfactory	
High school and more	6	8	14
High school only	4	10	14
Total	10	18	28

$r_{tet} = -.23$ (not significant); $\chi^2 = .156$.

4. Education: Shown in Tables 5, 6, and 7.

The performance rating criterion is shown in relation to each of the personal characteristics variables in Tables 2, 3, 4, 5, 6, and 7.

G. PROCEDURE AND RESULTS

On the basis of a knowledge of the concepts embodied in *AVA* profiles and a systematic analysis of the jobs involved, the general *AVA* score profile shapes were established by the company president as desirable applicant attributes and are shown in Figure 1.

The *AVA* Analyst (company president) rated the *AVA*'s of the 60 Managers and Assistant Managers in his employ at the time the standards were established.

By visually comparing these test profiles with the standards each was rated as either Favorable for success or Unfavorable for success on the basis of judged similarity or dissimilarity to this family of profile shapes.

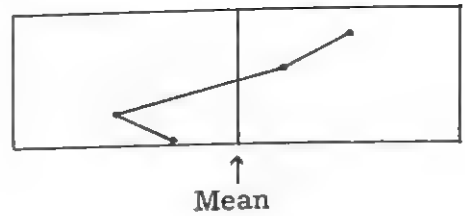
Table 8 reveals a significant relationship in the concurrent prediction situation for Office Managers.

Table 9 reveals a similar set of findings for the Assistant Office Managers hired before the introduction of *AVA* selection standards into the overall employment procedure.

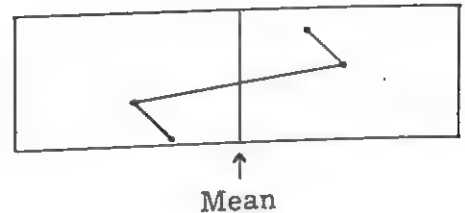
Subsequently, over a period of 10 to 11 months a total of 29 potential managers were employed with their *AVA*'s being rated in the same manner as the previous job incumbent samples. An attempt was made to preserve the screening considerations used before adoption of *AVA* in the selection process. Such factors as age, education, and kind of previous experience (most new Assistant Managers are hired without prior loan office experience) and interview judgments were used as usual with their restriction of range effects resulting. Some screening of the applicant group as a result of *AVA* profiles probably did enter into these selections; although these were believed to be restricted to the more extreme unfavorable *AVA* cases according to the company analyst. The results are shown in Table 10.

1. Best

Vector One
 " Two
 " Three
 " Four

2. Also Favorable

Vector One
 " Two
 ($r_{1.2} = +.93$) " Three
 " Four

3. Also Favorable

Vector One
 ($r_{1.3} = +.82$) " Two
 " Three
 " Four

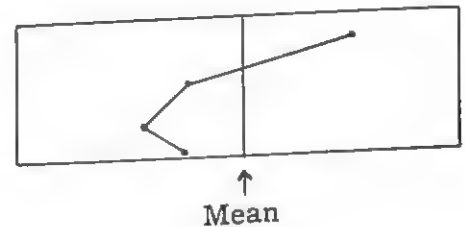


FIGURE 1
 AVA SCORE PROFILES USED AS RATING STANDARDS

TABLE 8
 PRESIDENT'S AVA RATINGS VS. SUPERVISORS' COMPOSITE PERFORMANCE RATINGS FOR
 OFFICE MANAGERS HIRED BEFORE AVA WAS USED

Performance rating	AVA ratings		Total
	Unfavorable	Favorable	
Satisfactory	7	16	23
Unsatisfactory	9	2	11
Total	16	18	34 = N

$r_{tet} = +.70$; $\chi^2 = 5.958$; $P = .018$ (corrected for continuity).

Table 10 reveals a potentially significant cross validation in that the P value of the χ^2 calculated indicates the obtained correlation value may represent a true value greater than zero, especially in view of some known restriction in the range of AVA data. While a P value of less than .05 was

TABLE 9

PRESIDENT'S *AVA* RATINGS VS. SUPERVISORS' COMPOSITE PERFORMANCE RATINGS FOR ASSISTANT OFFICE MANAGERS HIRED BEFORE *AVA* WAS USED

Performance rating	<i>AVA</i> ratings		Total
	Unfavorable	Favorable	
Satisfactory	4	6	10
Unsatisfactory	14	2	16
Total	18	8	26

$$r_{tet} = +.725; \chi^2 = 4.479; P = .04 \text{ (corrected for continuity).}$$

TABLE 10

PRESIDENT'S *AVA* RATINGS VS. SUPERVISOR'S COMPOSITE PERFORMANCE RATINGS FOR ASSISTANT OFFICE MANAGERS HIRED AFTER *AVA* WAS ADOPTED

Performance rating	<i>AVA</i> ratings		Total
	Unsatisfactory	Satisfactory	
Satisfactory	2	17	19
Unsatisfactory	5	5	10
Total	7	22	29

$$r_{tet} = +.67; \chi^2 = 3.628; P = .06 \text{ (corrected for continuity).}$$

desired, the tendency for the discrimination in the Table 10 data to occur sharply but only among the Satisfactory rated cases is a favorable as well as unusual feature for such selection data.

H. DISCUSSION OF RESULTS

As seen in Table 1, the sub criteria were extremely highly correlated with the composite criterion. The latter was used exclusively to test the *AVA* ratings for this reason. Although no direct reliability data are presented, the magnitude of these intercorrelations does serve to indicate that adequate criterion reliability probably does exist.

It would seem likely that the significant experience difference among Office Managers, as shown in Table 2, may be interpreted best in terms of more than two years being required to train an Office Manager adequately in this company. In view of the indication appearing in Table 3 that: the optimum age for satisfactory performance rating as a manager is 26 years and that a significant relation exists between progressively higher age status and less satisfactory performance ratings, the nature of this business seems to involve an emphasis on young, aggressively ambitious people. The data may well be too limited for such a sweeping interpretation. However, the known personality of the company's president, together with the temperament implications of the *AVA* profile selected as the most desirable, does coincide with

this interpretation. Although the comparisons of educational level and performance rating tend not to show significant relationships in Tables 5, 6, and 7, the near significant trend in Table 5 does parallel this tendency to stress ambition together with youth.

The significant validities of *AVA* profile ratings for the three sub groups of this study, shown in Tables 8, 9, and perhaps 10 as well, were produced by means of independent ratings of individual performances and *AVA*'s. As indicated in the description of geographical locations of the subjects, the individuals rated by district supervisors are widely scattered. Therefore, the performance ratings were of necessity the product of district supervisory judgment. In the case of the sub groups hired before the study was begun, it is likely that their known quality of performance had influence on the company president's determination of the desirable *AVA* profiles and the subsequent rating of their individual profiles against the standards finally specified. To the extent that this is true, the results for these Managers and Assistant Managers are in the nature of concurrent statistical representation of the *AVA* Profile segregation, achieved on the basis of both job appraisal and implicit employee appraisal.

However, the possible contamination of *AVA* Ratings with performance ratings cannot be in any way ascribed to the data involving those assistant managers hired with *AVA* under study. The difference from the first Assistant Manager group to the later Assistant Manager group of $r_{tet} = +.73$ to $r_{tet} = +.67$ seems a rather conservative amount of shrinkage in spite of the small numbers of cases involved. This, together with the weight of discrimination falling among the satisfactory performers, would seem to justify a tentative assertion of true validity and worthwhileness of further study.

I. SUMMARY

The *AVA* Profiles of three groups of Loan Office Managerial personnel were rated in terms of a hypothesized set of desirable profiles and these ratings compared with a composite district supervisory performance rating. Subject's age, experience, and education characteristics were also examined in terms of performance rating status. The results indicated significant age, experience, and *AVA* discriminations were possible.

J. CONCLUSIONS

1. The training or experience requirement for satisfactory performance in this loan office manager job seems to be approximately three years; at least something over two years.

2. The optimum age for this loan office manager job seems to be approximately 26 years with a tendency for higher age status to parallel lesser success.

3. The *AVA* profiles selected and the ratings of employee and applicant test-based profiles based on similarity to these standards seem to contribute to effective selection of potential office managers in this personal loan company.

4. Further follow-up verification of predictions of success on these and later employed individuals when they attain office manager status is seen as desirable and justified by these results.

REFERENCES

1. CLARKE, W. V. The construction of an industrial selection personality test. *J. of Psychol.*, 1956, **41**, 379-394.
2. McNEMAR, Q. *Psychological Statistics*. New York: Wiley, 1949. Pp. 344.

Walter V. Clarke, Inc.

500 Angell Street

Providence 6, Rhode Island

PERSONALITY PROFILE OF SELF-MADE COMPANY PRESIDENTS*

Walter V. Clarke Associates, Inc.

WALTER V. CLARKE

A. INTRODUCTION

In an early (1952) study of the validity of Activity Vector Analysis for differentiation among occupational samples, one of the average profiles of Activity Vector Analysis Scores derived was for successful business administrators and presidents using a selected sample of 57 persons. (The results of this study are cited in the *Manual for the Administration of Activity Vector Analysis*.)¹ This profile of scores was utilized later to hypothesize an ideal pattern for executive and presidential occupations as well as any others requiring a similar temperament syndrome. It was named the Administrative Pattern.

The purpose of this study was to verify the appropriateness or validity of this pattern for the occupational status of self-made business concern president. The stipulation of being a self-made company president introduced into our sampling procedure the requirement that each individual be known to have reached his position by means other than familial or political persuasion and inheritance of responsibility.

During the course of studies on a variety of job samples and the previously mentioned study of occupational differentiation by means of score profiles, a total of 14 so-called Standard Job Patterns had been derived and assigned deviation score values to constitute a full range of idealized temperament syndromes. Therefore, it was also decided to compare the distribution of the Standard Job Patterns (which includes the Administrative Pattern) which would be obtained in a sample of company presidents with the distribution which would be obtained in a diversely mixed occupational (but nonpresident) sample.

B. PROCEDURE

1. Test score profile cards were selected for presidents in the following manner: Account Executives and staff members of our organization were

*Received in the Editorial Office on February 1, 1956, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹Available on request—AVA Publications, Inc., Providence, R. I.

asked to name all self-made presidents with whom they had come in contact. No attempt was made to pick only those self-made presidents who were clients, there being many in our sample who are nonclients but who had voluntarily taken the test in the interest of learning about the test. A search of our files revealed a total of 154 test profiles for presidents named and known by our staff members.

No attempt at censoring, further selecting or cutting down of the sample was made. Thus the criteria for the selection of this sample was the naming of the men as self-made presidents by our staff members and the finding of the test score profiles in our files.

2. The mixed nonpresident occupational sample of test score profiles was selected as follows: Our files were systematically searched to find a wide variety of occupations, each of which could be represented by fifty cases wherever possible. Some occupations were included in this study with a lesser number, when an extensive search failed to reveal the desired number, in order to avoid unnecessary gaps in our sample. The occupational titles and numbers of cases resulting are shown in Table 1. This sample includes 1,067 persons and 23 occupations. (This is not to be confused with a sepa-

TABLE 1
COMPOSITION OF MIXED OCCUPATIONAL OR NON-EXECUTIVE POPULATION SAMPLE

	<i>N</i>
Machine Operators	50
Grinders	58
Engineers	56
Home Economists	31
Stenographers	38
Toolmakers	50
Insurance Agents	50
Public Health Nurses	50
Stewardesses	22
Forge Operators	20
Salesmen (General)	50
Draftsmen	50
Textile Overseers	48
Clerical Workers	52
Secretaries	49
Mill Workers	45
Electricians	50
Bank Tellers	48
Foremen	50
Textile Knitters	50
Textile Loopers	50
Accountants	50
Machinists	50
	1,067

rate sample of 1,067 cases which was drawn and used several years ago in connection with the construction of Activity Vector Analysis.)

3. A system of test score profile shape classification was applied to both samples. This system of shape classification includes fourteen standard patterns assigned idealized deviation score values and representing a range of shape intercorrelations from $r = +1.00$ through $r = -1.00$.

4. The distribution of shapes found in each of our samples was determined by tabulation and these distributions were compared. The differences between the distributions of shapes for the samples were tested by X^2 and a corresponding r_{tet} was calculated for the relationship between a profile shape dichotomy versus president-nonpresident occupational status.

5. Finally, a previously determined average score profile (the Administrative Pattern) for executive occupations was correlated with the profile score pattern for each of the 154 self-made presidents. These 154 correlation coefficients were converted into Fisher Z transformation values which were averaged in order to determine a corresponding mean correlation for the sample between the standard pattern and the existing patterns of profile scores in our sample.

C. RESULTS AND DISCUSSION

Table 2 presents the range of pattern shapes obtained in our samples in terms of per cent of each sample classified according to the idealized profile

TABLE 2
DISTRIBUTIONS OF STANDARD SCORE PROFILE SHAPES APPEARING IN A MIXED
OCCUPATIONAL SAMPLE AND A SAMPLE OF SELF-MADE PRESIDENTS

Standard pattern number	Mixed occupation sample	Self-made presidential sample
	4.78%	11.69%
1	10.78	3.25
2	5.53	1.95
3	1.12	3.90
4	6.65	34.42
5 (Administrative)	5.81	6.49
6	6.28	20.13
7	3.28	9.74
8	4.59	1.30
9	11.81	1.30
10	3.75	3.25
11	17.43	1.30
12	11.62	0.0
13	3.75	1.30
14	2.81	0.0
Unclassified		
	100.09%	100.02%
	$N = 1,067$	$N = 154$

TABLE 3
DISTRIBUTION OF SAMPLE SHAPES BY DESCENDING CORRELATION WITH ADMINISTRATIVE PATTERN

Standard pattern number	r with administrative pattern	Mixed sample	Per cent of self-made presidential sample
5 (Administrative)	$r = +1.00$	6.65%	34.42%
7	$+ .93$	6.28	20.13
6	$+ .90$	5.81	6.49
1	$+ .82$	4.78	11.69
4	$+ .73$	1.12	3.90
3	$+ .57$	5.53	1.95
8	$+ .41$	3.28	9.74
11	$+ .34$	3.75	3.25
2	.00	10.78	3.25
9	.00	4.59	1.30
10	$-.64$	11.81	1.30
14	$-.65$	3.75	1.30
13	$-.89$	11.62	0.0
12	$-.93$	17.43	1.30
Unclassified		2.81	0.0
		100.09%	100.02%

score shape most similar to each person's obtained score profile. It can be seen that several of these standard patterns in Table 2 show fairly substantial percentage differences between our samples. These include Nos. 5, 7, 12, 13, 10, 1, 2, and 8 in particular.

In order to clarify these differences the data in Table 2 are shown in Table 3 with the standard patterns listed in order of descending correlation with the Administrative Pattern (No. 5) which is of prime interest in this study.

An examination of Table 3 reveals a marked tendency for the majority of the cases in the Mixed Occupational Sample to occur with Standard Patterns less similar in shape to that of the Administrative Pattern, while the majority of the cases in the Self-Made Presidents Sample occur with Standard Patterns more similar in shape to that of the Administrative Pattern. Table 4 gives a rough summary of these trends.

The data shown in Table 4 were treated by a four-cell chi square calculation which revealed a value of 153.7 which is significant well beyond the .001 level of confidence. In addition a tetrachoric correlation coefficient was calculated which resulted in $r_{tet} = +.73$. It is apparent that these data strongly support the hypothesis that Activity Vector Analysis score profile shape is significantly related to self-made president occupational status as contrasted with other nonpresident occupational status.

TABLE 4
PERCENTAGES OF TWO SAMPLES WITH PROFILE SHAPES HAVING POSITIVE CORRELATIONS
AND ZERO OR NEGATIVE CORRELATIONS WITH THE ADMINISTRATIVE PATTERN

	Mixed non- presidential sample	Self-made presidential sample
Positively correlated	38.3% (N = 397)	91.6% (N = 141)
Zero and negatively correlated	61.7% (N = 640)	8.4% (N = 13)
	100.0% (N = 1,037)	100.0% (N = 154)

We next undertook to determine the degree of appropriateness or fit which characterizes our previously established administrative pattern with our sample of self-made presidents' actual patterns. Each individual's profile scores were converted into deviation values comparable to those used in our standard pattern. Thereafter, each individual president's profile was correlated with the standard profile.² Next, each correlation coefficient was transformed to Fisher Z transformation values which were then averaged.

TABLE 5
INDICES OF SIMILARITY OF PROFILE SHAPES FOUND IN SELF-MADE PRESIDENTS SAMPLE
WITH THE STANDARD ADMINISTRATIVE PROFILE SHAPE (N = 154)

Sample versus Standard			
Mean Fisher's z	=	+	1.3669
Standard Error of z	=	±	.0814
±1σ _z around Mean z	=	+	1.2855 through + 1.4483
Mean r	=	+	.878
±1σ _r around Mean r	=	+	.858 through + .895
(From σ _z limits)			

The mean z value was converted back to an r value as the best estimate of the mean correlation between sample profile shapes and the standard (Administrative) profile shape. The results are shown in Table 5.

The results shown in Table 5 show a very high relationship between our standard pattern derived in 1952 using 57 selected cases and the array of profile shapes appearing in our present sample as would be expected from the data shown in Tables 2, 3, and 4.

$$2r = \sqrt{\frac{(EX_1^2 + EX_{st}^2 - Ed^2)^2}{4(EX_1^2)(EX_{st}^2)}}$$

D. SUMMARY

The *AVA* profiles of 154 self-made company presidents were compared with a previously obtained average profile and found to be significantly similar to this reference profile. The distribution of pattern shapes classified in terms of 14 standard profiles was obtained for this sample and a general occupational sample. These distributions were compared and found to be significantly different.

E. CONCLUSIONS

1. A standard syndrome of temperament in the form of an Activity Vector Analysis score profile idealized shape for executive occupations has been shown to significantly characterize Self-Made President occupational status.
2. The range of Activity Vector Analysis score profile shapes has been shown to significantly parallel a range of occupational status to the extent of our sampling array of Self-Made President and Non-President occupations.

REFERENCES

1. CLARKE, W. V. Construction of an industrial selection personality test. *J. of Psychol.*, 1956, **41**, 379-394.
 2. ———. Manual for the Administration of Activity Vector Analysis. Providence, R. I.: *AVA Publications*, 1954.
 3. McNEMAR, Q. *Psychological Statistics*. New York: Wiley, 1949. Pp. 344.
- Walter V. Clarke Associates, Inc.*
500 Angell Street
Providence 6, Rhode Island

THE SYNONYM VOCABULARY TEST: STANDARDIZATION AND VALIDATION*

Department of Psychology, Pennsylvania State University

WILLIAM M. LEPLEY AND MARTIN L. ZEIGLER¹

A. PURPOSE

The purposes of this writing are: to review briefly the structure of the Synonym Vocabulary Test (1); to describe the standardization processes accomplished with a population of college freshmen; to present a summary of the results of these processes; and to present certain statistics as evidence of validity.

B. THE TEST

The rationale and construction of the test have been described in detail in a previous article (1). The following is quoted from the reference cited immediately above.

Though two individuals may be essentially equal in concept repertoire, they may differ significantly in language behavior variability; in the synonym vocabularies by means of which their concepts are expressed. It is proposed that such differences, if they exist, might reflect differences in adaptability or intelligence, as operationally defined by way of some criteria yet to be identified; and might do so with refinements of precision and meaning not yet achieved by other techniques for vocabulary assessment (1, 215-216).

For the processes to be described, the test was assembled in two forms, tentatively designated as Form I and Form II. Each form consisted of two parts: Part A and Part B. Each part contained eight, matching-type items. In each item the five stem-words were chosen for high frequency of use, as defined in *The Teacher's Word Book of 30,000 Words* (2). The synonym-response words, taken from a synonym dictionary (3), were also chosen on the basis of frequency of use, as defined above, and in such fashion as to attempt to maximize the range of difficulty within each item. Each item was furnished with 30 of these synonym-response words; thus providing an

*Received in the Editorial Office on February 2, 1956, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹This study was made possible by the cooperation of The Pennsylvania State University, Student Advisory Service, which is under the direction of Dr. Robert G. Bernreuter.

average of six correct matches for each stem-word. (To reduce matching by elimination the actual number was varied from four to eight.) In this format, each response word served as a correct match for one stem-word and as a mislead for the other four.

Within each part, in each form, the items were arranged in an order of difficulty, based on a tryout administration to 306 undergraduate students. These arrangements placed the least difficult items first in each of the orders. It should be emphasized that Form I and Form II contained the same two parts, consisting of the same items. The *only difference* between the two forms consisted in the order of the parts: Part *A* in Form I was identical with Part *B* in Form II; and Part *B* in Form I was identical with Part *A* in Form II. These reversed arrangements were made to counterbalance expected serial effects in the administration of this long testing task. The reader is reminded that each form required 480 responses on the part of the examinee.

C. THE STANDARDIZATION ADMINISTRATION

The test, in its two forms, was administered to 2,679 members of the freshman class at The Pennsylvania State University, during orientation week in 1954. The randomized distribution resulted in obtaining 1,343 cases with Form I and 1,336 cases with Form II.

The tests were administered without time limits. The time required varied from approximately 50 minutes to two hours. The proctors reported that essentially everyone completed the Parts *A* within one hour. Responses were recorded on answer sheets for machine scoring.

D. RESULTS

Reliability coefficients were computed for each of the two forms by correlating Part *A* scores with Part *B* scores. The resulting reliabilities were as follows: Form I: .87² ($N = 1,343$). Form II: .88 ($N = 1,336$). When these values were corrected by the Spearman-Brown formula they rose to .93 and .94 respectively.

As stated in a previous writing (1) the intent in dividing the total test into two parts for subsequent administration was based upon the forecast that a test of half-length might be useful as an instrument. In the reference cited, it was shown that the full-length test, *not* divided into parts, had a corrected, odd-even reliability of .97 for a mixed, college population of 306 students. In order to obtain comparable reliabilities from the data in the

²Throughout this report, all computations were based upon raw scores and ungrouped data.

present study, a 25 per cent, systematic sample (every fourth name in the alphabet) was drawn from the total population for each of the two forms. The Part *A* answer sheets for each form were segregated, rescored with odd and even keys, and the two pairs of arrays were used in the computation of odd-even reliabilities. These coefficients, when corrected by the Spearman-Brown formula were: for Form I, .92 ($N = 336$); and for Form II, .92 ($N = 334$). These reliability values appear to lend support to the proposal to use the two, half-length parts as equivalent forms.

Table 1 contains the summary of the results from the administration described above. As will be made clear later, the larger populations were subdivided to permit separate validity determinations. The liberal arts college and the engineering college categories need no further definition. The science category includes the colleges of chemistry and physics, mineral industries, and engineering. The non-science category includes the colleges of education, agriculture, home economics, business administration, liberal arts, and physical education. Thus, there is overlapping between the science category and the engineering category; and between the non-science category and the liberal arts category.

The values in Table 1 are largely self-explanatory. However, it might be pointed out that the anticipated serial effects appear to be confirmed by the consistently depressed means for the Part *B* scores in each population category. Furthermore, there appears to be some evidence that the eight items used in Part *B* of Form I and these same items used in Part *A* of Form II yield arrays of scores with slightly higher standard deviations than does the other set of eight common items administered as Part *A* in Form I and Part *B* of Form II.

E. VALIDATION AND COMPARATIVE STATISTICS

The freshman orientation week testing schedule at the Pennsylvania State University included the administration of an extensive battery of tests, which includes two measures singled out as especially relevant to the problems here investigated. One of these was the vocabulary section of the Pennsylvania College Placement Test in English.³ This standardized, unpublished test was developed locally to assist in the placing of freshmen into appropriate first courses and sections in English composition. The vocabulary section consists of 50, five-alternative, multiple-choice items. The other measure, relevant here, is the vocabulary section of The Pennsylvania State University

³This test was developed under the direction of Dr. J. S. Bowman, Professor of English Composition, who reports for it a reliability of .86; this based upon the correlation between alternate forms.

TABLE 1
SUMMARY OF POPULATION CHARACTERISTICS

Variables	Population categories									
	Liberal Arts (<i>N</i> = 147)		Non-Science (<i>N</i> = 649)		Engineering (<i>N</i> = 381)		Science (<i>N</i> = 537)		Combined* (<i>N</i> = 1,343)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Form IA	158.3	19.6	150.5	21.0	154.2	18.1	156.0	18.7	152.8	20.3
Form IB	153.5	24.1	144.5	25.7	150.6	22.1	152.6	22.6	147.5	24.4
Form I Total	311.8	42.2	295.0	44.9	304.9	38.7	308.6	40.0	300.3	43.2
	(<i>N</i> = 176)		(<i>N</i> = 679)		(<i>N</i> = 367)		(<i>N</i> = 536)		(<i>N</i> = 1,336)	
Form IIA	161.1	22.6	152.5	22.6	156.8	20.4	157.9	20.8	154.7	22.2
Form IIB	155.3	22.1	147.8	22.4	152.2	20.1	152.8	20.3	149.9	21.8
Form II Total	316.3	43.3	300.2	43.6	309.0	39.3	310.7	39.8	304.6	42.7

*The *N*'s in this table cannot be summed to obtain the combined populations reported here and foregoing for reasons explained in the text and also because some cases were lost when the data were assembled for the computation of the matrices from which the correlation coefficients in Table 2 were drawn for the special purposes of this paper. These losses were due to missing variables required by The Pennsylvania State University Student Advisory Service for elaborated studies of their own.

Academic Aptitude Examination.⁴ This standardized, unpublished, composite test, consisting of several sections, was developed for the express purpose of forecasting academic achievement in terms of the local grading system, which yields a summary value termed "grade point average."⁵ The vocabulary section of this composite examination consists of 120, five-alternative, multiple-choice items. The "grade point average," as defined in Footnote 5, for the *first semester only*, of the freshman year, was used as the criterion measure in the computations yielding the validity coefficients shown in the following section. As explained earlier in the text, the larger populations were subdivided in order to yield comparative values and some specificity of meaning to the various indices.

F. RESULTS

Table 2 contains the summarization of the validation information yielded by the procedures indicated for the populations described.

TABLE 2
SUMMARY VALIDATION INFORMATION: THE CRITERION IN EACH INSTANCE IS GRADE POINT AVERAGE FOR THE FIRST SEMESTER OF THE FRESHMAN YEAR

Predictor variables	Population Categories			
	Liberal Arts (N = 147)	Non-Science (N = 649)	Engineering (N = 381)	Science (N = 573)
SVT, Form I, Part A	.43	.41	.32	.37
SVT, Form I, Part B	.47	.42	.30	.37
SVT, Form I, Total	.46	.43	.32	.38
Engl. Pl. Voc.	.39	.32	.26	.31
Acad. Apt. Voc.	.40	.33	.27	.33
	(N = 176)	(N = 679)	(N = 367)	(N = 536)
SVT, Form II, Part A	.50	.42	.37	.38
SVT, Form II, Part B	.46	.41	.40	.40
SVT, Form II, Total	.49	.43	.39	.40
Engl. Pl. Voc.	.45	.36	.29	.32
Acad. Apt. Voc.	.45	.39	.35	.38

The values in Table 2 appear to justify the following face value interpretations: (a) In every instance, except one (wherein the values are equal), the academic aptitude, vocabulary predictor outperforms the English placement, vocabulary predictor. (b) In every comparison, except one (wherein

⁴The Pennsylvania State University Student Advisory Service reports a corrected, odd-even, reliability coefficient of .92 for this vocabulary section.

⁵At the time of this study, the grade designations were according to a six-point, numerical scale as follows: 3, 2, 1, 0, -1 and -2. The summary value, "grade point average," was obtained by multiplying the number of credits for each course by the grade earned, summing these products and dividing the sum by the total number of credits taken.

the values are equal), the Synonym Vocabulary Test predictors outperform those of the other two measures. (Though the differences referred to here are consistent in direction, they are not impressive in size and significance. The ratios of these differences to their standard errors range from 0.12 to 2.20.) (c) In every comparison except one (wherein the values are equal) all predictors have higher coefficients for the liberal arts and non-science categories than for the engineering and science categories. (d) When the Part *A* predictors (both forms) are compared with the Part *B* predictors (both forms) by transforming the coefficients to *Z*'s, computing means and transforming back to *r*'s, the two derived values turn out to be equal at .40. It thus appears that the depressed means for the Parts *B* scores are not associated with depressed predictive power.

Table 3 shows estimates of the communalities between the Synonym Vocabulary Test scores and the scores from the vocabulary section of the

TABLE 3
INTERCORRELATIONS BETWEEN SYNONYM VOCABULARY TEST SCORES AND THE OTHER
TWO VOCABULARY TEST VARIABLES

	Liberal Arts (<i>N</i> = 147)		Non-Science (<i>N</i> = 649)		Engineering (<i>N</i> = 381)		Science (<i>N</i> = 573)	
	Engl. Pl. Voc.	Acad. Apt. Voc.	Engl. Pl. Voc.	Acad. Apt. Voc.	Engl. Pl. Voc.	Acad. Apt. Voc.	Engl. Pl. Voc.	Acad. Apt. Voc.
SVT, I, A	.78	.87	.75	.79	.76	.78	.77	.80
SVT, I, B	.75	.81	.69	.73	.72	.78	.73	.80
SVT, I, Total	.78	.86	.74	.79	.77	.81	.77	.83
	(N = 176)		(N = 679)		(N = 367)		(N = 536)	
SVT, II, A	.81	.81	.73	.81	.74	.80	.76	.81
SVT, II, B	.74	.79	.67	.78	.73	.79	.73	.79
SVT, II, Total	.80	.83	.73	.82	.76	.82	.77	.83

Pennsylvania College Placement Test in English; and the vocabulary section of The Pennsylvania State University Academic Aptitude Examination, respectively.

The values in Table 3 show that the Synonym Vocabulary Test measures much in common with the other two measures. The coefficients vary from .67 to .87.

G. SUMMARY

The Synonym Vocabulary Test, in its two forms, as described herein, has been shown to have the following characteristics as it applies to predicting first semester, freshman, average grades: (a) Each of the half-length forms has an odd-even reliability of .92. (b) The instrument appears to show some

promise as a predictor of academic achievement. The validity in the populations studied ranges as follows: for liberal arts students, .43 to .50; for engineering students, .30 to .40; for non-science students, .41 to .43; and for science students, .37 to .40. These values, though ranging consistently somewhat higher, are not impressively larger than those from the other two vocabulary measures; and the intercorrelations between the sets of measures compared reveal rather high communalities. (c) Though the two, half-length forms yield similar values for means and standard deviations, some minor adjustments will need to be made in the construction of norm tables.

With regard to future development and use it is proposed to use the two half-length forms as equivalent.

REFERENCES

1. LEPLEY, W. M. The rationale, construction, and preliminary try-out of The Synonym Vocabulary Test. *J. of Psychol.*, 1955, **39**, 215-216.
2. THORNDIKE, E. L., & LORGE, I. The Teacher's Word Book of 30,000 Words. New York: Bureau of Publications, Teach. Coll., Columbia Univ., 1944.
3. Webster's Dictionary of Synonyms. Springfield, Mass.: Merriam, 1951.

Department of Psychology
Pennsylvania State University
University Park, Pennsylvania

REINFORCEMENT VALUE AS RELATED TO DECISION TIME*

Department of Psychology, Ohio State University

ERWIN J. LOTSOFF¹

A. INTRODUCTION AND PROBLEM

Resolution of conflict and decision making appear to be ubiquitous phenomena and permeate almost all behavior. Decision making has been treated by Festinger (3) and Cartwright (4) with regard to psychophysical phenomena, by Lewin (5, 6), Barker (1) and by Brown (2) and Miller (9) from a behavioristic position. Systematic treatment of the factors entering into decision making might aid in the prediction and control of human behavior.

This study is related to Barker's (1) but differs from it with regard to the methodology employed. Barker's Ss, 19 boys ages 9-11, were required to decide which one of two liquids they would drink in a "real" and "hypothetical" situation. Nine S's received the "real" situation first and after a 10-minute period were given the "hypothetical" situation; 10 S's received the reverse sequence. Prior to the actual experiment, each S ranked 7 liquids in order of preference. In the choice situations each liquid was paired with every other liquid. The present study differs from Barker's study in that it attempted to control the differences in the alternatives. This was accomplished by pretesting on a large group prior to the actual experiment. It was necessary to do this by virtue of the theory in which the problem was done. This study also utilized two separate groups for the "real" and "hypothetical" situations so that comparisons could be made between the groups. A larger sample was employed in this study as compared to Barker's, so that variability of performance might be reduced.

The problem has been formulated within Rotter's social learning theory of personality (10) and attempts to demonstrate the usefulness of decision time² as a measure of reinforcement value. Rotter's approach utilizes the

*Accepted for publication by Roy M. Dorcus of the Editorial Board, received in the Editorial Office on February 5, 1956, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹This paper is based upon a portion of a dissertation submitted to the Ohio State University in partial fulfillment of the requirements for the Ph.D. degree. The author wishes to express his appreciation to Dr. J. B. Rotter for his aid in this investigation.

²Decision time is defined as the time required to select one of two objects.

constructs of expectancy and reinforcement value for the prediction of behavior. Within the theory the potentiality of any behavior (B.P.) occurring is a function of (a) the expectancy (E) that the particular behavior will lead to a given reinforcement in a given situation and (b) the value of the given reinforcement (R.V.); $B.P. = f(E \text{ \& R.V.})$. Reinforcement value is inferred by the extent to which an S will select one object rather than another when expectancies for both are equal. Expectancy is the probability held by the individual that a behavior will lead to a given reinforcement and is measured for specific experiences for that behavior in that situation. Since behavior potential is a function of both expectancy and reinforcement, it would follow that if expectancy were held constant, behavior potential should be a function of reinforcement value. In this study expectancy was held constant since the reinforcement always occurred.

The study is methodological in that it attempts to ascertain the usefulness of decision time as a basis for measuring reinforcement value. In day-to-day experiences it has been observed that individuals frequently have to choose between one of two tasks; the tasks may sometimes be pleasant and sometimes unpleasant, and the time taken to decide varies quite considerably. There is another kind of choice situation which is often overlooked; this is the choice between doing something and doing nothing. Just as doing something has a reinforcement value so refraining from action has reinforcement value. Where doing nothing becomes more punishing with the passage of time than doing something, the decision or making the choice is reached rather rapidly. If the time taken to decide between two alternatives varied with the degree of "pleasantness" and "unpleasantness," it would be possible to use decision time as a measure of reinforcement value. Thus, it was hypothesized that the more punishing a pair of alternatives, with the differences between a given pair eliminated, the longer the decision time. It was also hypothesized that the greater the difference in reinforcement value between two alternatives the shorter the decision time.

The experiment was also designed to compare differences in decision time between choices followed by real reinforcement and choices followed by hypothetical reinforcement. When the S was given the reinforcement he chose, it was defined as *real*. When he was merely asked to state a choice verbally and to do nothing else, the reinforcement was defined as *hypothetical*. It was hypothesized that a reinforcement act in the hypothetical situation is not as strong as a reinforcement act in the real situation. Thus, for example, for a given choice situation two negative real choices would be more negative than two negative hypothetical choices and hence would require

a longer decision time. In the negative hypothetical situations the individual has a choice between doing nothing and doing something which, in this case, is selecting one of the liquids. It would appear that doing nothing becomes more punishing than doing something more quickly since in the negative hypothetical situation there is no real negative reinforcement to follow. Thus the negative hypothetical situation should take less time than the real.

Similarly it is hypothesized that the reinforcement value of the positive hypothetical situation is lower than the real situation. Thus, two positive real situations would be higher in reinforcement value than two positive hypothetical situations; consequently, there would be a shorter decision time when choices are made between the two real situations. In the positive hypothetical situation, the individual has a choice between doing nothing and doing something. Because doing nothing takes on negative consequences, he has to do something; but, since no actual positive reinforcement follows, the hypothetical positive choice takes more time than the real positive choice.

The study made it possible to determine the differences and similarities between behaviors followed by hypothetical or real reinforcement. In general, it might be expected that the principles derived from the real reinforcement acts can also be derived from the hypothetical reinforcement acts. If so, the more easily obtained hypothetical decisions might be useful in future experimental work.

B. METHOD

1. *Test Objects*

The names of 14 liquids were presented to 31 S's who were instructed to rank order them from most to least preferred. The resulting scale values were computed on the basis of mean ranks. The following pairs of liquids were selected for the experiments since both members of each pair had approximately equal means and standard deviations on the scale, pineapple and tomato juice, flat root beer and flat cola, a .02 per cent quinine solution and a 5 per cent acetic acid solution (vinegar).

2. *Materials*

Liquids used in this experiment tend to fall into three categories which vary from some degree of positiveness to negativeness. The quinine and vinegar solutions can rather easily be categorized as "strong negative." Flat root beer and pepsi-cola might be categorized as "weak negative." Many of the Ss' comments seemed to indicate that while these liquids did not taste good they were not as unpleasant as the quinine solution. A problem arises

in categorizing the pineapple and tomato juices. These liquids were at room temperature when presented to *Ss* and thus differed from the usual way in which they are served. While no *S* ever expressed marked disapproval, there were no indications that any *S* preferred to have his juices served at room temperature. As one *S* said, "These are good but not too good." Thus it is felt that perhaps the tomato-pineapple juice pair should be referred to as "weak positive" rather than "positive."

3. Subjects

Two groups of 33 *Ss* each were used, each *S* tested individually. The *real* group contained 17 males and 16 females; the *hypothetical* group contained 18 males and 15 females. All *Ss* were selected on a random basis from an alphabetical file containing all students taking the first course in elementary general psychology.

4. Procedure

One group was asked to choose and then to drink the selected liquid; the other group was asked merely to state verbal preferences. Each *S*, regardless of the group he was assigned to, tasted all six liquids which were presented in a random order and then he was asked to make choices between pairs of liquids. After testing each liquid, the *S* rinsed his mouth several times to avoid any after effects. Each liquid was presented in combination with every other liquid. The pairs of liquids were presented on cards. If the *S* was in the *hypothetical* group he was told: "I would like you to tell me which of the two liquids on this card you would prefer, if you had to drink one of them." In the *real* group, the *S* was told that after he selected the liquids he would have to drink approximately five cubic centimeters of it. A stop watch was used to measure decision time to tenths of a second.

C. RESULTS

Table 1 gives the rank order means and standard deviations of the liquids used in the choice situations. The difference in means between pineapple juice and tomato juice is 0.68, between flat pepsi-cola and flat root beer is 0.16, and between quinine and vinegar is 0.77. To determine whether decision time is a function of the inequality between the liquids placed in pairs, Tables 1 and 3 may be compared. Note that the most equal pair did not yield the longest decision time; nor did the pair with the greatest difference yield the shortest decision time. (Data in Table 1 are compared with 3 rather than 2 because both sets of choices were obtained in the hypothetical situation.) Thus the differences in decision time cannot be accounted for solely in terms of differences between the two members of the pair.

TABLE 1
RANK ORDER OF LIQUIDS HYPOTHETICALLY SELECTED FOR USE IN THE EXPERIMENT
($N = 31$)

Liquids	M	SD
Pineapple Juice	3.19	2.09
Tomato Juice	3.87	2.32
Flat Pepsi-Cola	8.03	2.02
Flat Root Beer	8.19	2.33
Quinine Solution	12.52	1.52
Vinegar Solution	13.29	1.08

The manner in which this experiment was carried out permits us to compare decision times for real and hypothetical situations of varying degrees of acceptability where alternatives were of differential value as reinforcements. Intra-comparisons and inter-comparisons of the groups used were made. These comparisons should answer questions as to whether it makes a difference if a reinforcement act is given realistically or symbolically and whether decision time is systematically related to the value of the reinforcement involved. The t test was used to compare the decision times of the groups.

For the *real* choices, the relationship of decision time to acceptability (reinforcement value) of the pair of alternatives when the within pair differences are controlled is given in Table 2. The decision times are all in the predicted direction. It takes less time to choose between two weak positive reinforcement acts than between two weak negative ones, and it takes less time to choose between two weak negative reinforcement acts than between two more negative ones. Differences in decision time for each of the three comparisons are significant at the .01 level of confidence.

TABLE 2
DECISION TIME FOR "REAL" CHOICES BETWEEN LIQUIDS: COMPARISON OF THE TWO WEAK POSITIVES, TWO WEAK NEGATIVES, AND TWO STRONG NEGATIVE CHOICES (IN SECONDS)

	M	SD	SEm
1. Pineapple Juice—Tomato Juice	1.88	.55	.09
2. Flat Pepsi-Cola—Flat Root Beer	2.64	.82	.14
3. Vinegar Solution—Quinine Solution	3.86	1.37	.24
Comparing 1 and 2.	$P < .01$		
Comparing 1 and 3.	$P < .01$		
Comparing 2 and 3.	$P < .01$		

For the *hypothetical* choices, the relationship of decision time to acceptability of the pair of alternatives when within pair differences are controlled is presented in Table 3. Decision times are all in the predicted direction and significant at the .01 level of confidence. The results are similar to those obtained with the *real* choices.

It was hypothesized that the greater the differences in reinforcement value between two alternative reinforcement acts the shorter the decision time. The difference between any one liquid of a pair and a liquid of an adjacent pair was given a step value of one. The difference between liquids in the extreme pairs was given a step value of two. When the liquids are within the same pair the step value is zero. The means in Table 4 were derived by averaging all possible comparisons for a given step value.

TABLE 3
DECISION TIME FOR "HYPOTHETICAL" CHOICES BETWEEN LIQUIDS (IN SECONDS):
COMPARISON OF THE TWO WEAK POSITIVES, TWO WEAK NEGATIVES,
AND TWO NEGATIVE CHOICES

	M	SD	SEm
1. Pinapple Juice—Tomato Juice	1.75	.45	.08
2. Flat Pepsi-Cola—Flat Root Beer	2.35	.58	.10
3. Vinegar Solution—Quinine Solution	3.09	.97	.17
Comparing 1 and 2. $P < .01$			
Comparing 1 and 3. $P < .01$			
Comparing 2 and 3. $P < .01$			

For the *real* situation, it can be seen in Table 4 that as the step value increases the decision time decreases; that is, an inverse relationship exists between them. In the *hypothetical* situation, differences between one and two-step, and between zero and two-step values are statistically significant. Although the zero-one step value difference is not significant, it is in the predicted direction.

TABLE 4
SIGNIFICANCE OF DIFFERENCES BETWEEN CHOICES MEASURED IN STEP VALUE
(IN SECONDS)

	M	N = 33 SD	SEm
<i>Real Situation</i>			
0. Zero step value	2.79	.83	.59
1. One step value	1.85	.13	.05
2. Two-step value	1.50	.12	.07
Comparing 0 to 1. $P < .01$			
Comparing 0 to 2. $P < .01$			
Comparing 1 to 2. $P < .01$			
<i>Hypothetical Situation</i>			
0. Zero step value	2.39	.56	.40
1. One step value	1.82	.10	.04
2. Two-step value	1.47	.09	.05
Comparing 0 to 1. $P < .20$			
Comparing 0 to 2. $P < .01$			
Comparing 1 to 2. $P < .01$			
0. Liquids in same category. 3 possible combinations			
1. Liquids in adjacent categories. 8 possible combinations			
2. Liquids in extreme categories. 4 possible combinations			

The conclusion appears warranted that it takes less time to decide between two reinforcements at opposite ends of a continuum than between two reinforcements which are close together on the same continuum. Decision time varies inversely with the degree of discrepancy between the value of two reinforcements.

Data in Table 5 indicate the relationship which exists between decision time and strength of a reinforcement when choices are presented *hypothetically* and *realistically*. When choices between two negatives are compared

TABLE 5

COMPARISON OF DECISION TIMES IN REAL AND HYPOTHETICAL SITUATIONS (IN SECONDS)

	Real M	Hypothetical M		
Quinine—Vinegar	3.86	3.09	$.05 > P < .10$	$P < .05$
Flat Pepsi-Cola—Flat Root Beer	2.64	2.35		$P < .10$
Tomato Juice—Pineapple Juice	1.88	1.73		$P < .20$

$P < .05$ and when the two weak negatives are compared $.05 > P < .10$. The differences between the two weak positives are not statistically significant. There were no differences between concrete and verbal behaviors, when the reinforcement acts could be characterized as weak positive, but there appear to be differences between concrete and verbal behaviors when choices had a negative value. The negative real reinforcement appears more punishing than the same negative reinforcement act presented symbolically.

D. DISCUSSION

These results tend to indicate that decision time might profitably be used to measure reinforcement value. Statistically significant differences were obtained in all tests except when the weak positives in the hypothetical and real situations were compared. With the weak positives the decision time for the hypothetical was less than for the real situation. That is to say, the real weak positive was a more punishing reinforcement act than the hypothetical weak positive. This finding raises an interesting question. Were the real weak positive choices actually less positive than the hypothetical choices? It should be recalled that the juices which the subject drank were at room temperature. Perhaps when these juices are warm they do not taste as pleasant as when they are cool. With the real weak positive choices we could have the following phenomenon taking place. When the S is told that he is going to taste some juices he has an expectancy and a minimal goal³ as

³The minimal goal is that point on a continuum where the occurrence of the act will increase the expectancy of the response but the next step down on the continuum will decrease the expectancy of the success. Reinforcement acts at the level of the minimal goal or higher would be positive, lower would be negative.

to how these juices should be; i.e., juice should be cold. In the real situation the *S* drinking the juice probably finds it below his minimal goal, while the *S* in the hypothetical situation does not experience this phenomenon. Thus, in the real positive situation, the *Ss* may have been receiving negative reinforcement acts which were more punishing than the hypothetical situations, consequently giving rise to longer decision times. In the hypothetical situation the *Ss* were not receiving the negative reinforcement acts. It might be conjectured that in this experiment the real choice was more punishing than the hypothetical for the weak positives.

Decision making has many implications for the clinical psychologist. One of the salient characteristics of maladjusted individuals is their rather marked inability to decide when faced with alternative behaviors as goals. A possible study which could be carried out might compare decision behaviors sampling various life situations before and after psychological therapy to determine to what extent the patient has "improved."

E. SUMMARY

This experiment was formulated within Rotter's social learning theory and the concept of reinforcement value was investigated. One group of *Ss* was asked to state verbal choices as to which one of a pair of liquids they would select if they had to drink one; another group of *Ss* was asked to state choices and then consume the liquid chosen. The time required to make the choice was obtained for the *real* and *hypothetical* groups. (a) Analysis of the data indicated that decision time varied inversely with the degree of discrepancy between the reinforcement values of the two acts. (b) Where pairs of negative reinforcement values of approximately equal strength are compared, decision time varies directly with the strength of the reinforcements. (c) Decision time for verbal preferences in this experiment appeared to be determined by the same principles as for choices between real objects. (d) There was a trend toward significance between verbal and actual choices which was consistent with the predictions made regarding their relative reinforcement values.

REFERENCES

1. BARKER, R. An experimental study of the resolution of conflict by children. In McNemar, Q., and Merrill, M. (Eds.) *Studies in Personality*. New York: McGraw-Hill, 1942. (Pp. 13-34.)
2. BROWN, J. S. The generalization of approach responses as a function of stimulus intensity and strength of motivation. *J. Comp. & Physiol. Psychol.*, 1942, **33**, 209-215.
3. CARTWRIGHT, D. A quantitative theory of decision. *Psychol. Rev.*, 1943, **50**, 595-602.

4. FESTINGER, L. Studies in decision: An empirical test of a quantitative theory of decisions. *J. Exp. Psychol.*, 1943, **32**, 411-417.
5. LEWIN, K. *Dynamic Theories of Personality*. New York: McGraw-Hill, 1935.
6. ———. The conceptual representation and the measurement of psychological forces. *Contrib. Psychol. The.*, 1938, **1**, No. 4.
7. LOTSOF, E. J. A methodological study of reinforcement value as related to decision time. Unpublished doctor's dissertation, Ohio State University, 1951.
8. MCNEMAR, Q. *Psychological Statistics*. New York: Wiley, 1949.
9. MILLER, N. E. Experimental studies on conflict. In J. McV. Hunt (Ed.) *Personality and Behavior Disorders*. New York: Ronald Press, 1944. (Vol. 1, pp. 431-465.)
10. ROTTER, J. B. *Social Learning and Clinical Psychology*. New York: Prentice-Hall, 1954.

Department of Psychology
University of California
Los Angeles 24, California

SUBLIMINAL PERCEPTION AND SUBCEPTION*

Department of Psychology, The Catholic University of America

JOSEPH H. VOOR¹

A. INTRODUCTION AND STATEMENT OF PROBLEM

1. *Introduction*

Recently there has been a great interest among psychologists in the effect of subliminal visual stimuli on conscious behavior. Though this area was investigated much earlier by a few psychologists (1, 23, 33), the current wave of interest was aroused by the experimental work on "perceptual defense." For an excellent summary of this work, the interested reader is referred to an article by Erikson (6).

The concept of perceptual defense, as proposed by Postman, Bruner, and McGinnies (30), and others (8, 27, 28), raises the thorny problem of how a person can defend himself against something that he doesn't see. Bruner and Postman (4) proposed two theoretical explanations of this phenomenon. The first explanation postulated a hierarchy of perceptual processes with some being veridical in nature, while others were largely affective responses. In addition to the veridical response, "other systematic reaction tendencies may be largely affective in nature and lead to various forms of avoidance responses" (4, p. 26). This view gives us two rather distinct types of perception—veridical perception and affective perception.

In accord with this theoretical position, Lazarus and McCleary (17, 18) sought to discover whether such discrimination without veridical awareness was possible. To overcome some of the criticism of McGinnies' experiment (22), they presented 10 nonsense syllables, instead of meaningful words, to their subjects. Five of these syllables were conditioned to produce a GSR. When these syllables were presented tachistoscopically, they found that the subjects responded with significantly greater GSR's to the shock syllables than to the neutral, i.e., non-shock, syllables, even though the subject could not identify the syllable shown. They interpreted these results as evidence of "subception"—"a process by which some kind of discrimination is made

*Received in the Editorial Office on February 6, 1956, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹This article was composed for the fulfillment of the publication requirement for the degree of Doctor of Philosophy in the School of Arts and Sciences of the Catholic University of America, Washington, D. C. The writer is indebted to Dr. W. D. Commins for the general supervision of this study.

when the subject is unable to make a correct conscious discrimination" (18). They felt that this sort of affective and autonomic response by which a person can discriminate between stimuli was sufficiently distinct from subliminal perception to warrant the new term. This study, subsequently confirmed (21, 31), supports the theory of two relatively independent types of perception. It suggests that there is an affective response to stimuli that is independent of the veridical response.

The second explanation of perceptual defense proposed by Bruner and Postman also postulated a hierarchy of perceptual responses. In this case, however, it was a hierarchy of thresholds within veridical perception. Such a hierarchy was postulated since it has been shown that there are "stages of perception" from a vague general outline to a clear perception of the stimulus (25). "In short, rather than thinking in terms of an 'absolute recognition threshold,' it is closer to the facts to posit a hierarchy of thresholds, each requiring a different quantity, and, perhaps, a different configuration of stimulation before a response is tripped off" (4, p. 27). According to this point of view there is only one type of perceptual process, the veridical process.

This theoretical position seems to be supported by the experimental work of Bricker and Chapanis (3), Murdock (26), and Howes (12). These authors maintain that the results of the Lazarus and McCleary experiment can be adequately explained in terms of the information received from the stimulus without postulating a subception process of autonomic discrimination without awareness. In two separate experiments (3, 26) these authors have demonstrated that even when the subject gives the wrong verbal response, he does obtain information from the presentation of the stimulus. They discovered that the subject could, by additional guesses, identify the missed syllable significantly better than chance. Since their presentation of the syllables was similar to that used by Lazarus and McCleary, they suggest that the subception effect can be explained in terms of this partial recognition of the shock syllable, and not in terms of any special autonomic discrimination mechanism. They feel that the most that can be said is that the GSR may be a more sensitive indicator of this partial recognition than other methods (11).

Though Bricker and Chapanis, Murdock, and Howes did not expressly relate their findings to the theoretical formulations of Bruner and Postman, their work appears to be based on the assumption that there is only the veridical type of perception with its various stages of clarity. Their data give some evidence that the perceptual process is a unitary one, insofar as

they have shown that the subject received "veridical information" from the stimuli, even when he made an incorrect verbal response. On the basis of this evidence, the authors maintain that the "subception" of Lazarus and McCleary is merely one stage of the ordinary veridical perception.

At the present time, therefore, both theoretical positions have some experimental support. The study of Lazarus and McCleary, however, did not include any measure of the information obtained from the syllable presentation beyond the rightness or wrongness of the subject's first response. It left unanswered the theoretical question whether the affective discrimination was made on the basis of the veridical information obtained from the stimuli. On the other hand, the studies by Bricker and Chapanis and Murdock demonstrate conclusively that information is obtained from incorrectly identified syllables. They fail, however, to substantiate their case against "subception," since they had no measure of the conditioned GSR.

2. *The Problem*

The problem of this study is, therefore, to investigate the relationship between the veridical information obtained from the stimuli and the GSR discrimination of the syllables. It attempts to answer the question: "Is the subception effect dependent on the veridical information obtained from the stimulus?" To throw some light on this question, the subception effect was studied at three subliminal levels of illumination. Thus it was possible to investigate the extent of the subception effect, and to determine whether the subception effect occurs even though the subject obtains no information, as indicated by the fact that he cannot identify the syllables above chance expectancy, from the presentation of the syllables. For this investigation the subception effect is defined in terms of finding significantly greater GSRs to the shock syllables than to the neutral syllables when the syllables are incorrectly identified. The term "subception" is used to indicate Lazarus and McCleary's interpretation of this effect as autonomic discrimination without awareness.

3. *Hypotheses*

On the basis of the previous research in this area by Bricker, Chapanis, and Murdock, two hypotheses were tested: (a) The subception effect would be obtained at the higher levels of illumination—those below the visual threshold but allowing recognition significantly above chance expectancy. (b) The subception effect would not be found when the verbal recognition of the syllables was not significantly above chance, i.e., when no partial veridical information was obtained from the stimulus.

B. THE EXPERIMENTAL PROCEDURE

To investigate the extent of the subception effect, the stimuli were presented at three subliminal levels of illumination. These levels were chosen so that the highest illumination gave about 50 per cent accuracy in the recognition of the stimuli, while the lowest illumination allowed only chance accuracy. A separate analysis of the subception effect was made at each of these three levels of illumination to determine the limits of the subception phenomenon.

The broader theoretical question concerning the interpretation of this phenomenon was investigated by the use of two measures of veridical information. One measure was objective; the other was subjective. The objective measure was the per cent correct responses made by the subject, while the subjective measure was the subject's degree of confidence in his response. Both of these measures were used in a study by Williams (33). The use of these measures made it possible to investigate the relationship between the subception effect and the amount of veridical information obtained.

1. *The Subjects*

For this study 16 subjects were used. They were undergraduate and graduate male students who volunteered their services. Since this experiment involved a perceptual task, a check was made to ascertain that their vision was approximately normal. Each subject was tested individually.

2. *Experimental Design*

The design of the experiment required the presentation of nonsense syllables on a beaded screen seven feet from the subject. The syllables were flashed on the screen one at a time by means of an automatic film strip projector to which a tachistoscopic shutter was attached. The size of the image on the screen was 2" by 12". The eight stimulus syllables were: HAZOX, ZEHAM, WEDYZ, WOZUM, MUWYD, MOYAW, XYMUW, HYWED. These syllables were put on film strip in order to assure equal clarity. The film strip contained 56 syllables. The order of the syllables was random except for the modification that each syllable appeared seven times. The same film strip was used over and over, but the subject was not aware of this, because of the number of syllables on the strip and the frequent interruptions in the presentation of the syllables.

The speed of exposure was constant for all subjects, but the illumination was varied according to the perceptual performance of each subject. The level of illumination at which the syllable was presented was controlled by

the *E* by means of a variable resistor connected in series with the projector lamp. The level of illumination at any given time was defined in terms of the setting on the variable resistor. Three different settings were used. The highest illumination setting was such that the subject was able to identify about 50 per cent of the presentations; the middle setting allowed between 25 per cent and 33 per cent correct choices; and the lowest setting gave approximately chance accuracy. Once these settings were determined for a subject, they remained fairly constant during the rest of the experiment.

Simultaneously with the presentation of the syllables, the *GSR* of the subject was recorded by means of a Lafayette Polygraph. Dry stainless steel electrodes $1\frac{1}{4}$ " square were attached to the palmar and dorsal surfaces of the *S*'s left hand. Before applying the electrode on the back of the hand, a small amount of electrode jelly was rubbed into the surface (2). The *GSR* was read as the maximum deflection of the pen which occurred within five seconds after the presentation of the syllable. The original measures in millimeters were converted to the log conductance, since this seems to be the preferred measure of changes in the *GSR* (9, 16) when further statistical analysis is required.

The testing was carried out in a semi-darkened room. The room was not soundproof, but most of the incidental noise of the equipment was masked by the constant noise of the motor used to drive the timing mechanism for the presentation of the stimuli. The presence of the noise did not appear to affect the subject's response. Though the subject could hear the equipment, he was not able to see it.

3. Testing Routine

a. Preliminary period. The subject was instructed that he was to identify the syllables that would be flashed on the screen. To aid him in this task he was given six cards, each with a different arrangement of the eight nonsense syllables. He was further told that whenever he saw a flash of light on the screen, he was to "guess" from his list of syllables, but he was not to use the same card for two successive presentations. This was introduced as a precaution against the subject "guessing" according to particular patterns of syllables. If he was fairly certain that he had seen the syllable, he was to name the syllable and add the number "1"; if he was doubtful, he was to guess a syllable and add "2" to indicate that this was a guess; if he saw no letters and had to make a blind guess, he was to pick out the first syllable that seemed likely and say the number "3" after it. He was then given 56 presentations, each syllable seven times, to determine the illumina-

TABLE 1
MEAN GSR TO NEUTRAL AND SHOCK SYLLABLES WHEN CORRECTLY IDENTIFIED

Subject		HVS		MVS		LVS	
		Neutral	Shock	Neutral	Shock	Neutral	Shock
1.	M	1.19	1.93	1.35	1.45	1.43	.00
	σ	.59	.18	.15	.28	.35	.00
	N	(12)	(7)	(30)	(8)	(7)	.00
		$p < .01$					
		$t = 3.89$					
2.	M	0.77	1.21	0.57	0.84	1.59	0.78
	σ	.65	.58	.46	.58	.10	.00
	N	(20)	(9)	(20)	(8)	(2)	(1)
		$p < .02$					
		$t = 2.71$					
3.	M	1.52	2.01	1.01	1.46	1.86	.00
	σ	.52	.26	.80	.86	.28	.00
	N	(30)	(19)	(15)	(8)	(2)	.00
		$p < .001$					
		$t = 4.45$					
4.	M	0.51	1.13	0.32	0.57	.00	0.99
	σ	.49	.55	.48	.60	.00	.10
	N	(15)	(22)	(6)	(17)	.00	(2)
		$p < .01$					
		$t = 3.50$					
5.	M	0.85	1.61	0.71	1.25	1.15	2.25
	σ	.47	.37	.58	.59	.00	.00
	N	(19)	(14)	(34)	(21)	(1)	(1)
		$p < .001$					
		$t = 5.03$					
6.	M	1.48	2.23	1.45	1.97	1.01	2.24
	σ	.57	.30	.56	.46	1.03	.09
	N	(12)	(10)	(10)	(10)	(4)	(3)
		$p < .001$					
		$t = 4.14$					
7.	M	0.40	1.57	0.87	1.27	0.39	0.52
	σ	.73	.76	.92	.73	.63	.79
	N	(30)	(26)	(14)	(5)	(7)	(11)
		$p < .001$					
		$t = 5.79$					
8.	M	0.72	1.32	0.63	1.29	0.70	1.04
	σ	.53	.38	.61	.70	.63	.06
	N	(18)	(18)	(13)	(6)	(7)	(2)
		$p < .01$					
		$t = 3.05$					

TABLE 1 (continued)

HVS		MVS		LVS	
Subject	Neutral	Shock	Neutral	Shock	Neutral
9. M	1.11	1.53	1.01	1.62	0.00
σ	.85	.65	.84	.94	**
N	(9)	(11)	(8)	(14)	(1)
10. M	0.92	1.17	0.80	1.37	0.48
σ	.79	.78	.76	.51	.63
N	(27)	(23)	(23)	(14)	(8)
					$p < .02$ $t = 2.65$
11. M	1.35	1.99	1.39	2.04	1.30
σ	.80	.17	.90	.25	.77
N	(21)	(8)	(7)	(11)	(13)
					$p < .05$ $t = 2.20$
12. M	0.77	1.39	0.68	1.20	0.69
σ	.50	.29	.67	.44	.56
N	(9)	(9)	(11)	(19)	(10)
					$p < .05$ $t = 2.17$
13. M	2.08	2.12	1.39	1.68	0.73
σ	.72	.71	1.16	1.11	1.03
N	(32)	(33)	(15)	(10)	(6)
14. M	1.29	1.44	0.86	1.36	1.63
σ	.55	.61	.70	.61	.18
N	(8)	(11)	(3)	(6)	(2)
					$p < .05$ $t = 2.17$
15. M	1.33	1.46	0.54	0.79	0.00
σ	.75	.65	.76	.65	**
N	(32)	(28)	(3)	(3)	(1)
16. M	1.37	1.39	1.42	1.78	1.48
σ	.79	.76	.68	.54	.06
N	(5)	(6)	(27)	(17)	(3)
Group mean	1.10	1.59	0.94	1.37	0.96
					1.23

this experiment the difference between the MIW_n (mean wrong GSR to neutral syllables) and the MIW_s (mean wrong GSR to the shock syllables) categories is the subception effect.

The data are treated separately for each of the three levels of illumination. For convenience, the high level of illumination, allowing about 50 per cent correct identification of the stimuli, will be referred to as the *HI'S* (high variac setting); the medium level, giving between 25 per cent and 33 per cent recognition, will be referred to as the *MI'S* (medium variac setting); and the low illumination level, allowing recognition just above chance expectancy, will be designated as the *LI'S* (low variac setting).

2. Results

Log conductance was used as the measure of change in the GSR. Table 1 gives the means and standard deviations of the GSR's for each of the 16 subjects, when they were able to identify the stimulus syllable. Turning our attention to the *HI'S* column, it is noted that 10 of the subjects show significant discrimination between the shock and neutral syllables. Of the remaining six subjects, Subject 10 shows significant discrimination at the *MVS*, but the other five manifest no significant difference in their GSRs to the shock and non-shock syllables. In all cases, however, the GSR was greater to the shock syllables than to the neutral syllables.

Table 1 also indicates a trend from greater discrimination at the *HVS* to less discrimination at the *LVS*. Of the 10 subjects who had significant differentiation at the *HVS*, only three maintained that degree of discrimination at the *MVS*. At the *LVS* only one subject could discriminate significantly between the shock and neutral syllables on the basis of his GSR. The same trend is noted in the mean GSRs of the total group given at the bottom of Table 1.

The relationship between this trend and the subject's degree of confidence in his verbal response is shown in Table 2. This table shows that at the

TABLE 2
THE DEGREE OF CONFIDENCE OF SUBJECTS IN THEIR CORRECT VERBAL RESPONSES AT
THREE ILLUMINATION LEVELS

Illumination	Per cent of certain responses ("1")	Per cent doubtful responses ("2")	Per cent guess responses ("3")
<i>HVS</i>	43%	42%	15%
<i>MVS</i>	18%	44%	32%
<i>LVS</i>	14%	31%	55%

HVS quite a high percentage of the correct responses were made with a great degree of confidence: 43 per cent "1" responses and 42 per cent "2" responses. At the *MIS* the percentage of "1" responses drops sharply, while the percentage of "3" responses increases. This indicates that, even though the subject does make the correct response, he is subjectively quite uncertain about it. Finally at the *LIS* the "3" responses predominate, thus pointing out an even greater lack of confidence.

Table 3 gives the means and standard deviations of the *GSRs* for each of the 16 subjects when they were unable to identify the syllable presented—the *MIV_n* and *MIV_s* categories. In only four cases do the differences reach statistical significance. Originally the plan was to investigate the subception effect on the basis of the individual records, but in view of the data, this was not feasible.

Corresponding to this lack of *GSR* discrimination, it was found that for the *MIV* categories the subjects made predominantly "2" and "3" responses. At the *HVS* about 35 per cent were "1" and "2" responses, while at the *LVS* about 95 per cent were "3" responses. Even though few of the subjects showed any significant discrimination, it should be noted that in the great majority of the cases, the *GSR* is greater to the shock syllables, even though these syllables were not correctly identified. This difference is evident at all three levels of illumination, but less so at the *MIS* and *LIS*, as is seen from the group means in Table 3.

In view of this difference between the *GSRs* to the shock and the neutral syllables, a further test was made using the combined data of all 16 subjects. For this test the *MIV_n* and the *MIV_s* scores of each subject, as recorded in Table 3, were used as the measure of *GSR* for that subject. This gave 16 pairs of scores. The significance of the difference between the *MIV_n* and the *MIV_s* based on these 16 scores, was determined by the "*t*" test for correlated means.

The mean *GSR* of the 16 subjects for the *MIV* categories is represented graphically in Figure 1. An inspection of the *MIV* columns of this figure indicates that the subception effect is present at all three levels of illumination. These differences between the *MIV_n* and *MIV_s* columns are statistically significant. The difference at the *HVS* gives a "*t*" of 4.12, which is significant at beyond the .002 level for 14 degrees of freedom; the difference of the means at the *MVS* is significant at beyond the .01 level of probability ("*t*" of 3.67); and the difference at the *LVS* yields a "*t*" of 2.28, which is significant at the .05 level of confidence.

It should be noted once again that there is a gradual weakening of the *GSR* discrimination from the *HVS* to the *LVS*. There was, however, significant discrimination at all three settings.

TABLE 3
MEAN GSR TO NEUTRAL AND SHOCK SYLLABLES WHEN INCORRECTLY IDENTIFIED— MW_n AND MW_s

Subject	HVS		MVS		LVS	
	Neutral	Shock	Neutral	Shock	Neutral	Shock
1.	M	1.27	1.51	1.53	1.27	1.38
	σ	.51	.32	.33	.49	.54
	N	(11)	(29)	(56)	(11)	(23)
2.	M	0.67	1.07	1.09	1.15	1.17
	σ	.64	.51	.61	.65	.79
	N	(11)	(24)	(33)	(5)	(9)
$p < .05$ $t = 2.03$						
3.	M	1.63	1.57	1.48	1.75	1.64
	σ	.69	.86	.71	.36	.79
	N	(22)	(28)	(38)	(7)	(7)
4.	M	0.35	1.12	0.70	0.40	1.06
	σ	.58	.65	.65	.57	.68
	N	(25)	(19)	(38)	(11)	(8)
$p < .001$ $t = 3.96$						
5.	M	1.01	1.22	0.72	0.81	0.77
	σ	.71	.57	.65	.53	.67
	N	(15)	(20)	(52)	(8)	(10)
6.	M	1.49	2.30	1.51	1.42	1.53
	σ	.64	.18	.72	.89	.76
	N	(10)	(10)	(20)	(19)	(19)
$p < .01$ $t = 3.66$						
7.	M	0.92	1.03	0.79	0.80	0.84
	σ	.89	.83	.84	.93	.91
	N	(19)	(21)	(15)	(30)	(38)
8.	M	0.83	1.12	0.96	0.68	1.06
	σ	.53	.49	.56	.49	.43
	N	(26)	(17)	(19)	(14)	(21)

TABLE 3 (continued)

Subject	HVS		MVS		LVS	
	Neutral	Shock	Neutral	Shock	Neutral	Shock
9.	M	0.83	0.94	1.12	0.45	0.67
	σ N (20)	.71 (20)	.80 (10)	.79 (20)	.73 (7)	.67 (6)
10.	M	0.84	1.01	0.97	0.84	0.91
	σ N (22)	.58 (22)	.55 (25)	.65 (27)	.69 (22)	.70 (27)
11.	M	0.96	1.38	1.15	1.07	0.99
	σ N (8)	.97 (8)	.71 (10)	.88 (20)	.97 (25)	.83 (42)
12.	M	1.22	1.11	0.83	0.89	1.04
	σ N (6)	.27 (6)	.64 (4)	.63 (25)	.56 (41)	.58 (35)
13	M	1.41	2.07	1.75	1.69	1.61
	σ N (28)	1.08 (28)	.76 (27)	1.01 (29)	.92 (19)	1.15 (33)
14.	M	0.98	0.94	1.40	0.87	0.99
	σ N (15)	.63 (15)	.76 (14)	.32 (11)	.74 (23)	.68 (18)
15.	M	0.75	1.25	1.23	0.96	1.39
	σ N (24)	.71 (24)	.88 (20)	.76 (8)	.88 (18)	.77 (14)
16.	M	1.44	2.29	1.22	1.41	1.30
	σ N (9)	.63 (9)	.05 (2)	.73 (40)	.57 (14)	.65 (19)
Group mean	1.04	1.37	1.07	1.18	1.03	1.15

$p < .02$
 $t = 2.58$

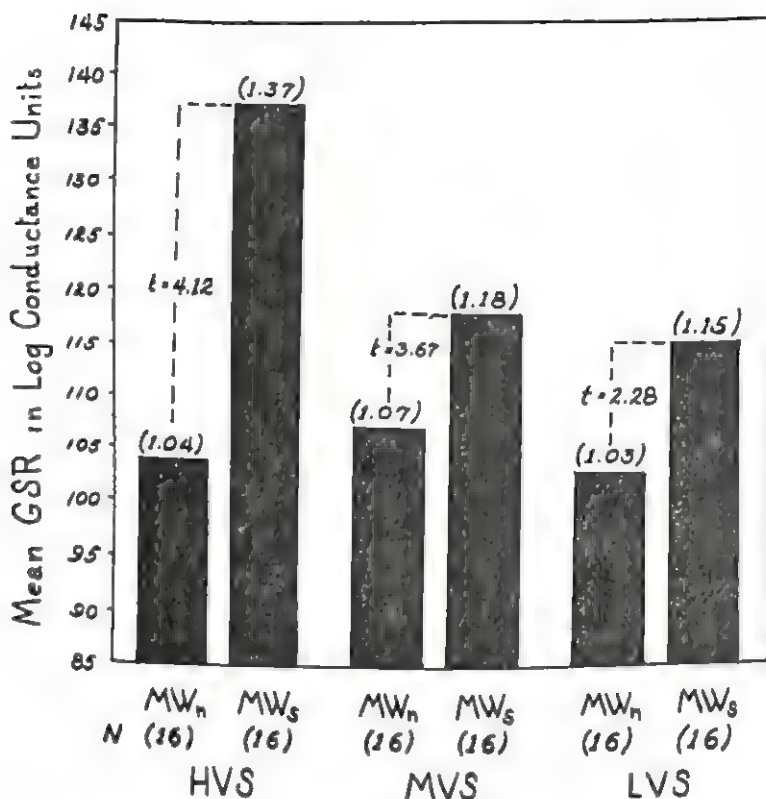


FIGURE 1
MEAN GSR OF 16 SUBJECTS (BASED ON THE MEAN GSR OF EACH SUBJECT) TO
INCORRECTLY IDENTIFIED SYLLABLES— MW_n AND MW_s —AT
THREE ILLUMINATION LEVELS

Since the statistical treatment of the data in this experiment differed somewhat from that used by Lazarus and McCleary in their study, it is rather hazardous to make direct comparisons. To obviate this difficulty, a second test of significance was made using the same method as in the original study. Their procedure combined the total number of responses of all 16 subjects and computed the significance of the difference on this total. Though this procedure is criticized by Jarrett and Henry (14) because it makes the unjustified assumption that all the subjects are from the same population, it was used here for the purpose of comparison.

The data of this test of significance are given in Figure 2. The "*t*" of 4.71 at the *HVS* is significant at beyond the .001 level of confidence; the "*t*" of 2.80 at the *MVS* is significant at beyond the .01 level; while the "*t*" of 1.94 at the *LVS* fails to reach statistical significance.

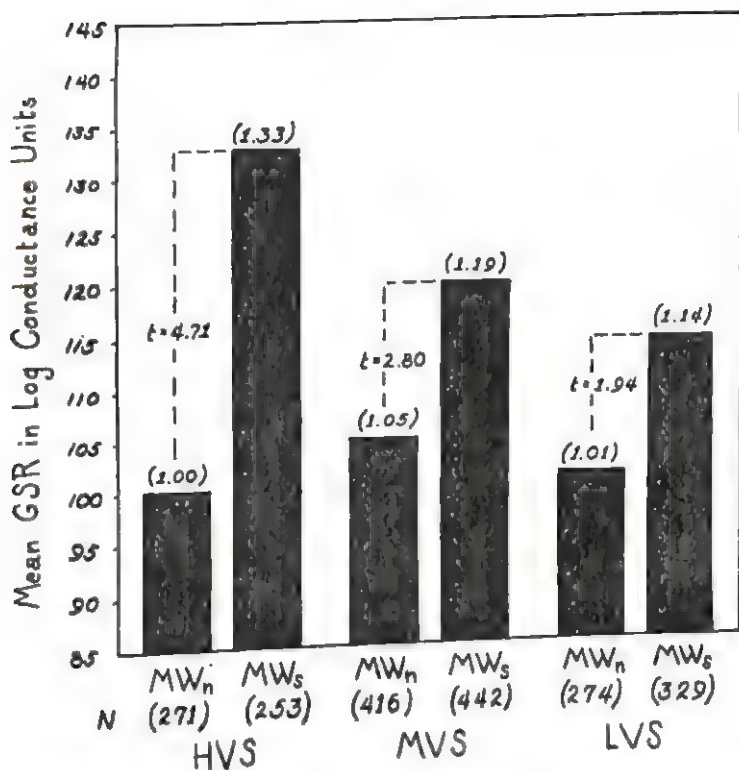


FIGURE 2
MEAN GSR OF 16 SUBJECTS (BASED ON TOTAL NUMBER OF RESPONSES) TO INCORRECTLY IDENTIFIED SYLLABLES— MW_n AND MW_s —AT THREE ILLUMINATION LEVELS

3. Response Frequency

Since it has been suggested by Murdock (26) that the subception effect may be due in part to a preference of the subjects for neutral syllables, the response frequency of neutral and shock syllables in this experiment was checked. The data for all the subjects are given in Table 4. From this

TABLE 4
RESPONSE FREQUENCY OF NEUTRAL AND SHOCK SYLLABLES WHEN THE RESPONSE IS INCORRECT

Stimulus syllable	HVS Per cent		MVS Per cent		LVS Per cent	
	WN	WS	WN	WS	WN	WS
Neutral	53 ($z = 1.98, p < .05$)	47	55 ($z = 4.10, p < .01$)	45	52	48
Shock	57 ($z = 4.45, p < .001$)	43	69 ($z = 15.9, p < .001$)	31	65 ($z = 10.9, p < .001$)	35

table it is apparent that the subjects in this experiment had a preference for using the neutral syllables when in doubt. Only in the case of the neutral syllables being presented at the *LVS* does this preference fail to reach statistical significance.

Since there exists such a decided preference toward a neutral verbal response, this may influence the accuracy of recognition of the neutral and shock syllables. It was pointed out that the subjects in the Lazarus and McCleary study were more accurate in their identification of the shock syllables (26), though the authors indicated that this was not true when a correction was made for the response preference of the subjects (18). The data in Table 5

TABLE 5
ACCURACY IN RECOGNITION OF NEUTRAL AND SHOCK SYLLABLES

Responses	<i>HVS</i>		<i>MVS</i>		<i>LVS</i>	
	Neutral	Shock	Neutral	Shock	Neutral	Shock
Total number	588	489	774	502	428	302
Per cent correct	51	52	31	36	17	18

give only slight evidence of a greater perceptual accuracy for the shock syllables in this study.

4. Discussion of Results

The data of this experiment strongly support the hypotheses under investigation. The first of these hypotheses was that the subception effect, as defined by Lazarus and McCleary, would occur at the higher illumination levels. This was verified by the fact that significantly greater *GSRs* were obtained for the shock syllables than for the neutral syllables, even though the syllables were not identified by the correct verbal response. This subception effect, or autonomic discrimination, occurred not only at the *HVS* and *MVS*, but also at the *LVS* as is shown by Figure 1.

The second and principal hypothesis of this study was that the subception effect would not be found when the verbal responses of the subject did not give accuracy above chance. In order to test the extent of the subception effect, three levels of illumination were used with the highest level allowing about 50 per cent accuracy in verbal response. The lowest level (*LVS*) was designed to give only chance recognition, but due to the fluctuating limen of the subjects during the testing period (33) and some errors on the part of the experimenter, it was found that this *LVS* allowed greater than chance accuracy. The use of the three levels of illumination, however, and the recording of the degree of confidence of the subjects in their verbal responses gave sufficient data to test this hypothesis.

The original interpretation of the subception effect as autonomic discrimination without awareness (18) has been criticized on the basis that the subject had a partial awareness of the syllable presented even though he made the incorrect verbal response (3, 12, 26). That the subject actually received information from these presentations was demonstrated by Bricker and Chapanis (3). Thus a measure of the information obtained from the syllable presentations at each illumination level was required in this study.

Since it was impractical to use the technique of Bricker and Chapanis in conjunction with the conditioning procedure of the original experiment, a different measure of information was needed. The measure chosen was that used in some early studies of subliminal perception by Williams (33) and Miller (23). They discovered that if a stimulus was presented at subliminal intensities, it would be identified significantly above chance, though there would also be many wrong responses. From this it was concluded that if the subject was getting any information from the presentation of a syllable at a given level of illumination, this should become evident by a better than chance recognition over a great number of exposures. For this experiment, therefore, the information threshold was defined as that level of illumination at which the subject can identify the syllables just above chance expectancy. If the subject cannot identify the syllables above chance accuracy this indicated that he received no veridical information from the presentations. As an additional index of the information obtained, the subject's degree of confidence in his responses was also recorded. This subjective measure of information was also based on the study by Williams (33).

Interpreting the data of this experiment on the basis of these two measures of information, it is noted that *GSR* discrimination was most significant when the most information was obtained from the presentation of the syllables, i.e., at the *HVS*. At this illumination level the subject was able to identify 50 per cent of the syllables, and showed a great degree of confidence in his responses: 43 per cent "1" responses (Table 2). As the illumination level was lowered, the subject manifested progressively less confidence in his response, even though it was correct. This lack of confidence is reflected in the *GSR* discrimination as well, as can be seen in Table 1. Thus it appears that there is a rather close relationship between the *GSR* discrimination of correctly identified syllables and the subject's confidence in his response.

On the basis of the great confidence and accuracy when the syllables are correctly identified at the *HVS*, it is justifiable to assume that the subject also gets information at this illumination when he chooses a wrong response. This assumption has been demonstrated by Bricker and Chapanis (3) and

Several recent studies (3, 12, 26) have criticized this interpretation offered by Lazarus and McCleary. These studies have all pointed out the fact that the results explained by subception could be interpreted on the basis of the information that the subject received even when he gave a wrong verbal response. There is no need to postulate a special autonomic discrimination mechanism. In terms of the nature of the perceptual process, these authors seem to take the position that there is only veridical perception. They admit, however, that this process involves various stages. The present study was designed to test the validity of this criticism, and thereby to shed some light on the theoretical problem of the nature of the perceptual process.

The experiment was patterned after the original experiment of Lazarus and McCleary. A larger number of presentations and a larger number of subjects, 16, were used, so the data could be analyzed separately for the three illumination levels. Also two measures of the information obtained from the stimulus presentations were introduced. The one measure was a subjective report of the subject's confidence in his verbal response, while the other was an objective measure based on the percentage of correct responses made at a given illumination setting. The three illumination settings were such that the high setting gave 50 per cent correct responses while the low setting allowed only 17 per cent recognition.

It was hypothesized that if the subception effect was due to the information obtained from the stimulus presentations, this effect should not be found when the subject could not identify the syllable above chance expectancy by his verbal report. Though the low illumination setting was found to give significantly greater verbal accuracy than expected by chance and the subception effect was significant at this setting, a definite relationship was evident between the amount of information the subject received from the stimuli and his ability to discriminate on the basis of the conditioned GSR. As the GSR discrimination was highly significant, $p = .002$, at the high illumination setting but barely significant, $p = .05$, at the low setting, this was interpreted as support for the hypothesis of this study. Additional evidence for this interpretation was found in the fact that the subjects had a significant preference for using neutral responses when in doubt. The results of this experiment, therefore, corroborated the criticism made by Bricker and Chapanis (3), Murdock (26), and Howes (12).

On the basis of the data of this experiment the following conclusions may be drawn with some justification.

1. There exists a definite relationship between the information received from a stimulus presentation and the ability to discriminate on the basis of the GSR.

2. When no information is gotten from the presentation, indicated by verbal recognition below chance expectancy, the subception effect also fails to reach statistical significance.

3. There is no evidence of autonomic discrimination *without* awareness. Rather the GSR seems to be mediated by the partial recognition that the subject gets from the presentation.

4. In regard to the theoretical problem of the nature of the perceptual process, the data of this experiment supports the position that there is only veridical perception which involves various stages of clarity.

REFERENCES

1. BAKER, L. The influence of subliminal stimuli upon verbal behavior. *J. Exp. Psychol.*, 1947, **20**, 84-99.
2. BITTERMAN, M. E., KRAUSKOPF, J., & HOLTZMAN, W. The galvanic skin response following artificial reduction of basal resistance. *J. Comp. & Physiol. Psychol.*, 1954, **47**, 230-234.
3. BRICKER, P., & CHAPANIS, A. Do incorrectly perceived tachistoscopic stimuli convey some information? *Psychol. Rev.*, 1953, **60**, 181.
4. BRUNER, J., & POSTMAN, L. Perception, cognition, and behavior. *J. Personal.*, 1949-50, **18**, 14-30.
5. COOK, S. W., & HARRIS, R. E. The verbal conditioning of the galvanic skin response. *J. Exp. Psychol.*, 1937, **21**, 202-210.
6. ERIKSON, C. W. The case for perceptual defense. *Psychol. Rev.*, 1954, **61**, 175-182.
7. FREEMAN, J. T. Set or perceptual defense. *J. Exp. Psychol.*, 1954, **48**, 283-288.
8. GILCHRIST, J. C., & NESBERG, L. S. Need and perceptual change in need-related objects. *J. Exp. Psychol.*, 1952, **44**, 369-377.
9. HAGGARD, E. A. Experimental studies in affective processes: II. On the quantification and evaluation of measured changes in skin resistance. *J. Exp. Psychol.*, 1945, **35**, 46-56.
10. ———. On the application of analysis of variance to GSR data: I. The selection of an appropriate measure. *J. Exp. Psychol.*, 1949, **39**, 361.
11. HOWES, D. H., & SOLOMON, R. L. A note on McGinnies' "Emotionality and perceptual defense." *Psychol. Rev.*, 1950, **57**, 229-235.
12. HOWES, D. A statistical theory of the phenomenon of subception. *Psychol. Rev.*, 1954, **61**, 98-110.
13. HUMPHREY, L. G. Extinction of PGR following two conditions of reinforcement. *J. Exp. Psychol.*, 1940, **27**, 71.
14. JARRETT, R. F., & HENRY, F. M. The relative influence on error of replicating measurements or individuals. *J. of Psychol.*, 1951, **31**, 175-180.
15. LACEY, O., & SIEGEL, P. An analysis of the unit of measurement of the galvanic skin response. *J. Exp. Psychol.*, 1949, **39**, 122.
16. LACY, O., LEWINGER, N., & ADAMSON, J. Foreknowledge as a factor affecting perceptual defense and alertness. *J. Exp. Psychol.*, 1953, **45**, 169-174.
17. LAZARUS, R., & MCCLEARY, R. A. Autonomic discrimination without awareness: an interim report. *J. Personal.*, 1949, **18**, 171-179.
18. ———. Autonomic discrimination without awareness: A study of subception. *Psychol. Rev.*, 1951, **58**, 113-122.

19. LAZARUS, R. Is there a mechanism of perceptual defense? A reply to Postman, Bronson, and Gropper. *J. Abn. & Soc. Psychol.*, 1954, **49**, 396-398.
20. LONGENECKER, E., KRAUSKOPF, J., & BITTERMAN, M. Extinction following alternating and random partial reinforcement. *Amer. J. Psychol.*, 1952, **65**, 580-587.
21. LOWENFIELD, J. Verbal inhibition in subception. Paper read at East. Psychol. Ass., Boston, 1953.
22. MCGINNIES, E. Emotionality and perceptual defense. *Psychol. Rev.*, 1949, **56**, 244-251.
23. MILLER, J. G. Discrimination without awareness. *Amer. J. Psychol.*, 1939, **52**, 562-578.
24. MOELLER, G. The CS-UCS interval in GSR conditioning. *J. Exp. Psychol.*, 1954, **48**, 162-166.
25. MOORE, T. V. Process of Abstraction. *University of California Publications in Psychology*, 1910, **1**, 73-197.
26. MURDOCK, B., JR. Perceptual defense and threshold measurements. *J. Personal.*, 1954, **22**, 565-571.
27. POSTMAN, L., BRONSON, W., & GROPPER, C. L. Is there a mechanism of perceptual defense? *J. Abn. & Soc. Psychol.*, 1953, **48**, 215-225.
28. POSTMAN, L., & BROWN, D. The perceptual consequences of success and failure. *J. Abn. & Soc. Psychol.*, 1952, **47**, 213-225.
29. POSTMAN, L., & BRUNER, J. Perception under stress. *Psychol. Rev.*, 1948, **55**, 314-324.
30. POSTMAN, L., BRUNER, J., & MCGINNIES, E. Personal values as selective factors in perception. *J. Abn. & Soc. Psychol.*, 1948, **43**, 142-154.
31. RUBENFIELD, S. Stimulus generalization in subception. Paper read at East. Psychol. Ass., Boston, 1953.
32. SCHLOSBERG, H., & STANLEY, W. A simple test of normality of 24 distributions of electrical skin resistance. *Science*, 1953, **117**, 35-37.
33. WILLIAMS, A. C. Perception of subliminal visual stimuli. *J. of Psychol.*, 1938, **6**, 187-199.

Department of Psychology
Bellarmine College
2000 Norris Place
Louisville 5, Kentucky

EFFECTS OF ESTHETIC SURROUNDINGS: II. PROLONGED AND REPEATED EXPERIENCE IN A "BEAUTIFUL" AND AN "UGLY" ROOM*

Department of Psychology, Brandeis University

NORBETT L. MINTZ¹

A. THE PROBLEM

A previous paper by Maslow and Mintz (1) reported that when Ss spent 10-15 minutes rating a series of face-photographs, Ss tested in a "beautiful" room rated the faces as having significantly more "energy" and "well-being" than Ss tested in either an "average" or an "ugly" room. In discussing these results, one of their questions was: "Were these merely short-term effects; would the Ss adapt to the rooms with time and negate the initial differences obtained?" (1, p. 252). The differences might simply reflect either activity appropriate to a "laboratory" situation, or initial adjustments to the room conditions. If this were so, then Ss having prolonged or repeated experience in less "experimental" circumstances either would show no effects at all, or effects that would rapidly diminish with time. This report utilizes material, obtained from the examiners of the previous report, that pertains to this problem.

B. METHOD

The "beautiful" ("B") room and the "ugly" ("U") room referred to in this report were the ones fully described previously (1, p. 247). The "B" room was pleasantly decorated and furnished to give the appearance of an attractive, comfortable study; the "U" room was arranged to appear as an unsightly storeroom in a disheveled, unkept state. Controls for odor, noise, light, time of day, etc., helped restrict differences to visual-esthetic qualities.

The Brandeis undergraduates who were the male and the female examiners referred to in the previous report were also the "subjects" of this study. They were told that they were to be examiners in an experiment "on facial stereotypy" (1, p. 249). The examiners thus did not know that they were

*Received in the Editorial Office on February 7, 1956, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

¹The writer extends his appreciation to the Maslows for inspiring this research, to R. M. Held, J. B. Klee, and R. B. Morant for helpful discussion, and especially to "Sheila" and "Sid" for their cooperation. A Brandeis University research fellowship provided the necessary leisure time.

testing the effects of esthetic surroundings, and were *unaware that they were to be "subjects" themselves*. Therefore, the examiners' behavior can show what happens when people are not acting as *Ss* in an "experiment."

The examiners each spent six sessions (two per week) testing a total of 32 *Ss*. They tested *Ss* concurrently; while one examiner tested someone in the "B" room, the other examiner was testing another *S* in the "U" room. The first week had two sessions, on successive days, and each was an hour long; the second week had two sessions, separated by one day, and each was two hours long; the third week had two sessions, on successive days, and each was an hour long. The examiners spent the whole of one session in the same room; they would switch rooms on alternate sessions. Each examiner thus spent three sessions in the "B" room and three sessions in the "U" room.

Two measures plus observational notes form the basis of this report. The first measure was the same six-point, two-dimension rating scale used to test the effects of conditions upon *Ss* of the first report. The dimensions rated were "energy/fatigue" and "displeasure/well-being" for a series of 10 negative print face-photographs. Each face was rated as being very, rather, or slightly "fatigued," or slightly, rather, or very "energetic"; likewise for the dimension "displeasure/well-being." Each judgment had a weight of from 1-6; a total score was computed by summing the 10 rating-weights for each dimension and averaging the two sums. This average, which could range from 10-60, then represented a score of "energy" and "well-being."² Before the formal part of the experiment began, the writer tested each examiner (as if they were *Ss*) for 15 minutes in one of the two rooms followed by a 15-minute retest with duplicate photographs in the other room. This "practice" period served both to show the examiners the procedure, and to obtain data from these examiners that would be comparable to data obtained from the 32 *Ss* they would test. The examiners also administered the rating scale to themselves at the end of each session. Each examiner thus had three self-tests in the "B" room and three in the "U" room. This procedure enabled weekly checks on any effects. The purpose given for these retests was "to establish the reliability of the tests."

The effects of conditions upon the examiners' general interest in and enjoyment of the testing situation might be inferred from a comparison of testing-times for the 32 pairs of test situations. Since the examiners tested *Ss* concurrently and since the test procedure and the time each pair of *Ss* entered the rooms was identical for both examiners, the number of times

²See (1) for more complete details and samples of the negative prints.

an examiner in the "U" room finished before an examiner in the "B" room can provide a second measure of the effects. Besides recording testing-times, the writer took observational notes whenever possible.

The data will be examined as follows: In these more natural circumstances, will *any* effects be found? Will any effects that are found *rapidly decrease* after initial adjustments (*i.e.*, after the first week)?

C. RESULTS OF RATING SCALE

Table 1 shows that there were differences between scores for both the 15-minute "practice" periods and the prolonged sessions. The two examiners had higher scores (more "energy" and "well-being") in the "B" room. An analysis of variance for two subjects with repeated measures was computed to test these differences between prolonged sessions. The results are presented in Table 2. The over-all *F* test for the difference between sessions

TABLE 1
RATING SCALE TOTAL SCORES FOR THE "B" ROOM AND THE "U" ROOM

Rooms*	Examiners as Ss "Practice" periods		Prolonged sessions, Examiners testing Ss					
	"B"	"U"	First week "B"	"U"	Second week "B"	"U"	Third week "B"	"U"
"Sheila"	39.0	36.5	35.0	33.5	37.0	33.0	36.0	33.5
"Sid"	38.5	35.0	36.0	34.0	38.0	32.0	36.0	32.5
Means	38.75	35.75	35.50	33.75	37.50	32.50	36.00	33.00
Difference	3.00		1.75		5.00		3.00	

*"Sheila" was in the "B" room for the first "practice" period and the "U" room for the second; "Sid" was in the "U" room for the first "practice" period and the "B" room for the second. "Sheila" was in the "U" room for the first session of each week and the "B" room for the second; "Sid" was in the "B" room for the first session and the "U" room for the second.

TABLE 2
ANALYSIS OF VARIANCE FOR SCORES IN PROLONGED "B" AND "U" ROOM SESSIONS

Source	df	Mean Square	F	p
Sessions	5	7.5	18.75	<.01
Examiners	1	.1	—	
E X S	5	.4		
Total	11			

was significant well beyond the .01 level of confidence. The difference between examiners was negligible. Table 3 presents the analysis of the mean difference in scores obtained between a "B" room and "U" room session of the same week. Scores were significantly higher in the "B" room for *each* of the three weeks of prolonged sessions. Table 3 also shows that the mean difference in scores for the first week was the *smallest* obtained during the

TABLE 3
MEAN DIFFERENCE BETWEEN "B" AND "U" SESSIONS OF THE SAME WEEK

First week			Second week			Third week		
$\bar{X}\Delta$	t^*	p	$\bar{X}\Delta$	t^*	p	$\bar{X}\Delta$	t^*	p
1.75	2.74	<.05	5.00	7.81	<.001	3.00	4.76	<.01

*Based upon the E X S mean square, with 5 *df*, of Table 2.

three weeks, while that for the *second* week was the largest. Thus, the effects did *not* decrease after the first week.

D. RESULTS OF TESTING-TIME COMPARISONS

Table 4 tabulates which examiner finished testing first for each of the 32 pairs of test situations. When "Sheila" was in the "U" room and "Sid" was in the "B" room, "Sheila" finished before "Sid" a total of 13 times, compared with twice when she was in the "B" room and "Sid" was in the "U" room. When "Sid" was in the "U" room and "Sheila" was in the "B" room, "Sid" finished before "Sheila" a total of 14 times, compared with three times when he was in the "B" room and "Sheila" was in the "U"

TABLE 4
NUMBER OF TIMES AN EXAMINER ENDS FIRST IN THE "B" ROOM AND IN THE "U" ROOM

Weeks Testing rooms	One		Two		Three		Totals		p
	"B"	"U"	"B"	"U"	"B"	"U"	"B"	"U"	
"Sheila" first	0	3	2	7	0	3	2	13	.02
"Sid" first	1	4	1	6	1	4	3	14	.01
Combined sums	1	7	3	13	1	7	5	27	<.001

room. The total combined sum shows that the examiner in the "U" room finished testing before the examiner in the "B" room 27 times out of the 32 situations; this difference was significant beyond the .001 level of confidence. The combined sums in Table 4 show that the examiner in the "U" room finished before the examiner in the "B" room seven out of eight times the first week, 13 out of 16 times the second week, and again seven out of eight times the third week. Thus, once again the effects did *not* decrease after the first week.

E. OBSERVATIONAL NOTES

The following are the notes taken by the writer. It must be emphasized that they are *selective* notes; that is, the writer recorded only those comments and behavior that appeared relevant.

"Sheila"—During the 15-minute "practice" period in the "B" room,

she handled and admired an ashtray and remarked enthusiastically upon a piece of sculpture. She commented, "The rug on the floor doesn't quite match the rest of the décor." When taken to the "U" room for the second 15-minute period, she remarked, while rating a picture, "They all look more *blah* (fatigued) in here." As she was led to the "U" room for the first testing session, she exclaimed, "I have to start in here?" After testing two Ss she asked the writer, "Can't we change rooms now, or something to break up the monotony? I'm falling asleep in there." On the second session of the first week "Sheila" tested four Ss in the "B" room; when told she was through for that day she said, "Only four subjects? I was just getting warmed up."

"Sheila" spent the first two-hour session of the second week in the "U" room. After the first hour, while the examiners had a five-minute break, she smiled, stretched, and said (to "Sid"), "The dungeon is all yours." She assumed that they were to switch after one hour. When she was told this was not the case, she asked, "You mean I've got to spend the next hour in there too?" She was told that she must test four more Ss in the "U" room. When the next S arrived, "Sheila" led him down the corridor in the direction of the "U" room. The corridor is a dead end, at the end of which, on the left wall, is the door opening into the "U" room. However, instead of opening this left-wall door, she turned to the *right* wall and opened the only other door at the end of that corridor, mistakingly leading her S into the women's toilet! Has the rat in his *T* maze ever performed a more classical avoidance response?

After testing this S in the correct room, "Sheila" approached the writer and told him she was tired and was developing a headache. "Could we quit early?" She was told that this "would ruin the experimental design." She unhappily returned to the "U" room. The writer found that while waiting for this next S, "Sheila" had fallen asleep.

The second two-hour session "Sheila" knew she was to spend in the "B" room. As she met the writer, she remarked, "A two-hour session today? Good; I really feel like working tonight." This mood was sustained throughout that session. The first session of the third week she again was in the "U" room. Throughout this session she waited in the corridor for her Ss; previously she had waited for her next S in whichever room she was testing. At the end of this session she asked, "Tomorrow is the last day (of the experiment)? How nice; I'll end testing upstairs (the "B" room)."

"Sid"—During the 15-minute "practice" period in the "U" room "Sid" remarked, as he entered the room, "Ugh, what the hell did they do, empty

the whole building's junk in here?" When taken to the "B" room for the second 15-minute period, he asked, "Is this M's office? Pretty nice. He really makes things comfortable for himself." At the end of the first session, which he spent in the "B" room, he said, "I was really beginning to feel like a wheel in here, sitting in a swivel chair and making like a psychologist. It's a lot of fun." When he was told, at the start of the second session, that he was to test in the downstairs ("U") room (the rooms were simply called the "upstairs" and the "downstairs" room throughout the experiment) "Sid" remarked, "I knew I couldn't be in 'heaven' forever." After testing one S he approached the writer and said, "I think it's pretty stupid to use this room for an experiment." He was reminded that we were replicating a previous experiment and "must do it exactly the same." At the end of the second session he said, in a question-assertion manner, "Next week I return upstairs, eh?"

The first two-hour session of the second week "Sid" was in the "B" room. He gave a pleased smile when the writer told "Sheila" that they would not switch rooms after the break. The second two-hour session he spent in the "U" room. At the break he was generally aggressive in his conversation. He complained about having to add up the scoresheets (a procedure initiated the first session of the second week), and in a half-jocular, half-aggressive manner told the writer, "I think I'll just let you add them up from now on." When "Sheila" asked for a match, "Sid" looked in his shirt pockets, found a pack of cigarettes, but no matches. After the break, as he led his next S toward the "U" room, he took a cigarette from his shirt pocket and lit it from matches he suddenly found in his pants pocket. At the end of that session, he asked, "Next week I go back upstairs?" The last session of the third week, "Sid" commented as he entered the "U" room, "Well, this is the last time I'll have to see this hole."³

F. DISCUSSION

The results of the rating scale, the testing-time comparisons, and the observational notes all demonstrated the significant effects of esthetic surroundings. Furthermore, these effects were not limited to initial adjustments. In fact, there were indications on the rating scale and in the observational notes that the two-hour sessions of the second week *exaggerated* the effects.

³On completing this study, the writer told the examiners the real purpose of the experiment. They evidenced surprise at the whole procedure. Especially impressive was their reaction to the notes; they were not aware that their activities were in such close relationship to the room conditions, though they both realized that they did not prefer to test in the downstairs ("U") room.

May we infer that if the groups previously reported had prolonged experience in these conditions, they too would have continued to show differences in effects? It may be remembered that the 15-minute "practice" periods obtained data from the examiners comparable to data of the previous report. Therefore, the scores of the *initial* 15-minute "practice" period that "Sheila" spent in the "B" room and that "Sid" spent in the "U" room may be compared to the means reported previously for the "B" room and "U" room groups. The means for the groups were 37.99 and 31.81 respectively, with a standard deviation of 4.8 (1, p. 251). "Sheila" had an initial "B" score of 39.0 and "Sid" had an initial "U" score of 35.0 (see Table 1). Although their scores were slightly higher, they were well within one standard deviation for the respective group means; in other words, their scores were not significantly different from scores obtained in similar conditions from unselected college students.

Since the examiners did not have atypical results, it might be expected that the "B" room group would have continued to show higher scores. If bias was introduced by the choice of examiners, it possibly was reflected in the observational notes. That is, other examiners may not have had such gross behavioral changes, or may not have been as free in expressing their feelings.

In their report Maslow and Mintz also asked, "Were the Ss affected by the rooms *per se*? . . . the results could have been obtained via the effect of conditions upon the examiners" (1, p. 252). In like manner it may be asked, "Were the effects upon the examiners brought about by the rooms *per se*?" It is conceivable, for example, that the examiner in the "U" room finished testing first simply because his S "hurried" the procedure along. However, there is ample evidence from the behavioral notes to indicate that the results were not solely determined by the Ss being tested. There probably was a complex relationship whereby the esthetic conditions affected the Ss and the examiners, and the Ss and examiners in turn affected each other.

G. SUMMARY AND CONCLUSIONS

The present study investigated whether the effects of esthetic surroundings reported previously simply reflect either "laboratory" activity or initial adjustments to the room conditions. During a period of three weeks, two examiners, *unaware* that they were "subjects" for this study, each spent prolonged sessions testing Ss in a "beautiful" room and in an "ugly" room. On a rating scale, the examiners had short-term effects similar to those reported previously; furthermore, during the entire three weeks of prolonged sessions the ratings continued to be significantly higher in the "B" room. The test-

ing-time comparisons showed that an examiner in the "U" room usually finished testing more quickly than an examiner in the "B" room. Observational notes showed that in the "U" room the examiners had such reactions as monotony, fatigue, headache, sleep, discontent, irritability, hostility, and avoidance of the room; while in the "B" room they had feelings of comfort, pleasure, enjoyment, importance, energy, and a desire to continue their activity. It is concluded that visual-esthetic surroundings (as represented by the "B" room and "U" room) can have significant effects upon persons exposed to them. These effects are not limited either to "laboratory" situations or to initial adjustments, but can be found under naturalistic circumstances of considerable duration.

REFERENCE

1. MASLOW, A. H., & MINTZ, N. L. Effects of esthetic surroundings: I. Initial short-term effects of three esthetic conditions upon perceiving "energy" and "well-being" in faces. *J. of Psychol.*, 1956, **41**, 247-254.

*Department of Psychology
Brandeis University
Waltham, Massachusetts*

